

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2006, 13:32:22 ; Search time 4710 Seconds
(without alignments)
11187.660 Million cell updates/sec

Title: US-10-613-990A-31
Perfect score: 927
Sequence: 1 gtcattgtacgcgatcatg.....gtaactttcttactgggttt 927

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_scs.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	927	100.0	927	6	CQ760542 Sequence
2	927	100.0	39675	1	Continuation (47 o
3	927	100.0	338534	1	EC00093 Escherich
4	920.6	99.3	28445	1	AE005174_55
5	920.6	99.3	98450	1	BA000007_54
6	912.6	98.4	86898	1	AE016772 Escherich
7	912	98.4	1468	1	ECODYE M10044 E.coli dye
8	910.2	98.2	107203	1	AE005674_45
9	910.2	98.2	225944	1	AE016993 Shigella
10	715.6	77.2	88037	1	AL627284 Salmonell
11	715.6	77.2	284233	1	AE016849 Salmonell
12	714	77.0	13852	1	AE008916 Salmonell
13	714	77.0	55700	1	AE017220_47
14	712.4	76.9	85229	1	CP000026_45
15	615.6	66.4	822	1	AY832936 Escherich
16	612.6	66.1	619	1	AY832937 Escherich
17	610.8	65.9	614	1	AY832938 Escherich
18	610.8	65.9	614	1	AY832940 Escherich

19	607.2	65.5	612	1	AY832939	Escherich
20	577.2	62.3	726	6	AR385083	Sequence
21	562.4	60.7	564	1	AP267619	Escherich
22	562.4	60.7	564	1	AP267620	Escherich
23	562.4	60.7	564	1	AP267621	Escherich
24	560.8	60.5	564	1	AP267627	Escherich
25	560.8	60.5	564	1	AP267628	Escherich
26	560.8	60.5	564	1	AP267629	Escherich
27	560.8	60.5	564	1	AP267631	Escherich
28	560.8	60.5	564	1	AP267632	Escherich
29	560.8	60.5	564	1	AP267633	Escherich
30	560.8	60.5	564	1	AP267635	Escherich
31	560.8	60.5	564	1	AP267637	Escherich
32	559.2	60.3	564	1	AP267622	Escherich
33	559.2	60.3	564	1	AP267623	Escherich
34	559.2	60.3	564	1	AP267624	Escherich
35	559.2	60.3	564	1	AP267625	Escherich
36	559.2	60.3	564	1	AP267626	Escherich
37	559.2	60.3	564	1	AP267630	Escherich
38	559.2	60.3	564	1	AP267636	Escherich
39	559.2	60.3	564	1	AP267638	Escherich
40	557.6	60.2	564	1	AP267634	Escherich
41	549	59.2	549	1	AY698108	Escherich
42	547.8	59.1	551	1	AY698120	Shigella
43	546.8	59.0	550	1	AY698088	Shigella
44	546.8	59.0	550	1	AY698102	Shigella
45	545.8	58.9	549	1	AY698114	Escherich

ALIGNMENTS

RESULT 1
CQ760542
LOCUS CQ760542 927 bp DNA linear PAT 03-MAR-2004
DEFINITION Sequence 31 from Patent EP1382686.
ACCESSION CQ760542
VERSION CQ760542.1 GI:44904047
KEYWORDS Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.

REFERENCE 1
AUTHORS Ishikawa,Y., Imaizumi,A., Matsui,K. and Kojima,H.
TITLE Method for producing target substance by fermentation using a bacterial strain lacking the Arca gene
JOURNAL Patent: EP 1382686-A 31 21-JAN-2004;
Ajinomoto Co., Inc. (JP)
FEATURES
Location/Qualifiers
1..927
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
101..817
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAP32793.1"
/db_xref="GI:44904048"
/translation="MQTPHILIVDELVTRLKSIPEAGYVFEATDGAEMHQLS
EYDNLVIMDINLPKGLLARELREQANVALMPLTGRDNEVDKILGLETGADYIT
KPPNPRELTIRARNLLSRNLTGTVSERRSVSYKFGWELDINSRLIGPDGRQYK
LPREFFAMLHFCENPGKIQSRAELLKQMTGRELKPHDRTVDVTIRIRKHFESPTD
PELIATIHGEYRFGCDLED"

ORIGIN
Query Match 100.0%; Score 927; DB 6; Length 927;
Best Local Similarity 100.0%; Pred. No. 3.8e-262;
Matches 927; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCATTGTACGCCGATCATGTTAATTCAGATGCATCAGGTCAGGACTTTGCT 60
|||||

```
Db      1  GTCATGTTACGCCGATCATGTTAAATTTGCAGCATCATCAGCAGGTGAGGACTTTTGT 60
Qy      61  ACTTCTGTTTCGATTTAGTTGGCAATTTAGTAGCAAAACATGCAAGACCCGCAATTCT 120
      |||
Db      61  ACTTCTGTTTCGATTTAGTTGGCAATTTAGTAGCAAAACATGCAAGACCCGCAATTCT 120
Qy     121  TATCGTTGAAGACGAGTTGGTAACACGCAACACGTTGAAAAGTATTTTCGAAGCGGAAG 180
      |||
Db     121  TATCGTTGAAGACGAGTTGGTAACACGCAACACGTTGAAAAGTATTTTCGAAGCGGAAG 180
Qy     181  CTATGATGTTTTCGAAGCGACAGATGGCGCGGAAATGCATCAGATCCTCTCTGAATATGA 240
      |||
Db     181  CTATGATGTTTTCGAAGCGACAGATGGCGCGGAAATGCATCAGATCCTCTCTGAATATGA 240
Qy     241  CATCAACCTGGTGATCATGGATATCAATCTGCCGGGTGAAGACGGTCTTCTGTAGCGGG 300
      |||
Db     241  CATCAACCTGGTGATCATGGATATCAATCTGCCGGGTGAAGACGGTCTTCTGTAGCGGG 300
Qy     301  TGAACCTGGCGAGCAGGCGAATGTTGCGTTGATGTTCTGACTGCCGCGTGACAAAGAA 360
      |||
Db     301  TGAACCTGGCGAGCAGGCGAATGTTGCGTTGATGTTCTGACTGCCGCGTGACAAAGAA 360
Qy     361  CGATAAAATTTCTCGGCTCGAAATCGGTGCAGATGACTACATCACCACCAACCGTTCAACCC 420
      |||
Db     361  CGATAAAATTTCTCGGCTCGAAATCGGTGCAGATGACTACATCACCACCAACCGTTCAACCC 420
Qy     421  GGTGAACTGACGATCTGTGCGAAGCACTTACTGTCCCGTACATGAAATCTGGGTACTGT 480
      |||
Db     421  GGTGAACTGACGATCTGTGCGAAGCACTTACTGTCCCGTACATGAAATCTGGGTACTGT 480
Qy     481  CAGCGAAGACGTCGTAGCGTTGAAAGCTACAAAGTTCAATGTTGGAAGCTGGACATCAA 540
      |||
Db     481  CAGCGAAGACGTCGTAGCGTTGAAAGCTACAAAGTTCAATGTTGGAAGCTGGACATCAA 540
Qy     541  CAGCGGCTTCGTTGATCGGCCCTGATGGCGAGCAGTACAAGCTGCCGCGCAGGTTCCG 600
      |||
Db     541  CAGCGGCTTCGTTGATCGGCCCTGATGGCGAGCAGTACAAGCTGCCGCGCAGGTTCCG 600
Qy     601  CGGCATGCTTCACTTCTGTGAAAACCCAGCGAAAATTCAGTCCCGTGTGAATCTGTA 660
      |||
Db     601  CGGCATGCTTCACTTCTGTGAAAACCCAGCGAAAATTCAGTCCCGTGTGAATCTGTA 660
Qy     661  GAAATGACCGCGGTGAGCTGAAACCGCACCGCTACTGTAGACGTGACGATCCGCG 720
      |||
Db     661  GAAATGACCGCGGTGAGCTGAAACCGCACCGCTACTGTAGACGTGACGATCCGCG 720
Qy     721  TATTCGTAAACATTTTGGAACTCTACGCCGATACGCCGGAATTCATCGCCACCATTCACGG 780
      |||
Db     721  TATTCGTAAACATTTTGGAACTCTACGCCGATACGCCGGAATTCATCGCCACCATTCACGG 780
Qy     781  TGAAGGTTATCGTTCGCGGTGATCTGGAAGATTAAATCGGCTTTACGACCGTCAAAAA 840
      |||
Db     781  TGAAGGTTATCGTTCGCGGTGATCTGGAAGATTAAATCGGCTTTACGACCGTCAAAAA 840
Qy     841  AACGGCGCTTTTATAGCGCGTTTTATTTTCAACCTTATTTCCAGATACGTAATCATC 900
      |||
Db     841  AACGGCGCTTTTATAGCGCGTTTTATTTTCAACCTTATTTCCAGATACGTAATCATC 900
Qy     901  GTCCGTTGTAACTTCTTTACTGGCTTT 927
      |||
Db     901  GTCCGTTGTAACTTCTTTACTGGCTTT 927
```

RESULT 2
U00096_46/c
WPCOMMENT
Sequence split into 47 fragments LOCUS U00096 Accession U00096
Fragment Name Begin End
U00096_00 1 110000
U00096_01 100001 210000
U00096_02 200001 310000
U00096_03 300001 410000
U00096_04 400001 510000

```
500001 610000  
500001 710000  
700001 810000  
800001 910000  
900001 1010000  
1000001 1110000  
1100001 1210000  
1200001 1310000  
1300001 1410000  
1400001 1510000  
1500001 1610000  
1600001 1710000  
1700001 1810000  
1800001 1910000  
1900001 2010000  
2000001 2110000  
2100001 2210000  
2200001 2310000  
2300001 2410000  
2400001 2510000  
2500001 2610000  
2600001 2710000  
2700001 2810000  
2800001 2910000  
2900001 3010000  
3000001 3110000  
3100001 3210000  
3200001 3310000  
3300001 3410000  
3400001 3510000  
3500001 3610000  
3600001 3710000  
3700001 3810000  
3800001 3910000  
3900001 4010000  
4000001 4110000  
4100001 4210000  
4200001 4310000  
4300001 4410000  
4400001 4510000  
4500001 4610000  
4600001 4639675  
Continuation (47 of 47) of U00096 from base 4600001 (U00096 Escherichia coli K-12 MG1655)

Query Match 100.0%; Score 927; DB 1; Length 39675;  
Best Local Similarity 100.0%; Pred. No. 7.9e-262;  
Matches 927; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Qy 1 GTCATGTTACGCCGATCATGTTAAATTTGCAGCATCATCAGCAGGTGAGGACTTTTGT 60  
Db 38429 GTCATGTTACGCCGATCATGTTAAATTTGCAGCATCATCAGCAGGTGAGGACTTTTGT 38370  
Qy 61 ACTTCTGTTTCGATTTAGTTGGCAATTTAGTAGCAAAACATGCAAGACCCGCAATTCT 120  
Db 38369 ACTTCTGTTTCGATTTAGTTGGCAATTTAGTAGCAAAACATGCAAGACCCGCAATTCT 38310  
Qy 121 TATCGTTGAAGACGAGTTGGTAACACGCAACACGTTGAAAAGTATTTTCGAAGCGGAAG 180  
Db 38309 TATCGTTGAAGACGAGTTGGTAACACGCAACACGTTGAAAAGTATTTTCGAAGCGGAAG 38250  
Qy 181 CTATGATGTTTTCGAAGCGACAGATGGCGCGGAAATGCATCAGATCCTCTCTGAATATGA 240  
Db 38249 CTATGATGTTTTCGAAGCGACAGATGGCGCGGAAATGCATCAGATCCTCTCTGAATATGA 38190  
Qy 241 CATCAACCTGGTGATCATGGATATCAATCTGCCGGGTGAAGACGGTCTTCTGTAGCGGG 300  
Db 38189 CATCAACCTGGTGATCATGGATATCAATCTGCCGGGTGAAGACGGTCTTCTGTAGCGGG 38130  
Qy 301 TGAACCTGGCGAGCAGGCGAATGTTGCGTTGATGTTCTGACTGCCGCGTGACAAAGAA 360  
Db 38129 TGAACCTGGCGAGCAGGCGAATGTTGCGTTGATGTTCTGACTGCCGCGTGACAAAGAA 38070  
Qy 361 CGATAAAATTTCTCGGCTTCGAAATCGGTGCAGATGACTACATCACCACCAACCGTTCAACCC 420


```

Db 38069 CGATAAAATTCCTCGCCTCGAAATCGGTGCAGATGACTACATCACCAACCGTTCAACCC 38010
QY 421 GCGTGAACCTGACGATTTCGTGCACGCAACTACTGTCCTGATCCGATCATGAATCTGGGTACTGT 480
Db 38009 GCGTGAACCTGACGATTTCGTGCACGCAACTACTGTCCTGATCCGATCATGAATCTGGGTACTGT 37950
QY 481 CAGCGAAGAACCTGCTAGCGTTGAAAGCTCAAGTTCATGTTTCAATGTTGGGAATCGACATCAA 540
Db 37949 CAGCGAAGAACCTGCTAGCGTTGAAAGCTCAAGTTCATGTTTCAATGTTGGGAATCGACATCAA 37890
QY 541 CAGCGTTTCGTTGATCGGCCCTGATGCGGAGCAGTACAGCTGCGCGGAGCGAGTTCGG 600
Db 37889 CAGCGTTTCGTTGATCGGCCCTGATGCGGAGCAGTACAGCTGCGCGGAGCGAGTTCGG 37830
QY 601 CCGCATGCTTCACTTCTGTGAAAAACCCAGGCAAAATTCAGTCCGCTGCTGAACCTGCTCAA 660
Db 37829 CCGCATGCTTCACTTCTGTGAAAAACCCAGGCAAAATTCAGTCCGCTGCTGAACCTGCTCAA 37770
QY 661 GAAATGACCGCGCGCTGAGCTGAAACCGCAGCAGCTACTGTGAGACGTGACGATCCGCG 720
Db 37769 GAAATGACCGCGCGCTGAGCTGAAACCGCAGCAGCTACTGTGAGACGTGACGATCCGCG 37710
QY 721 TATTCGTAACATTTTCGATCTACGCGGATACGCGGAGTACGCGGAAATCATCGCACCATTCACGG 780
Db 37709 TATTCGTAACATTTTCGATCTACGCGGATACGCGGAGTACGCGGAAATCATCGCACCATTCACGG 37650
QY 781 TCAAGGTTATCGCTTCTGCGGTGATCTGAAGATTAAATCGGCTTACACCGTCAAAAA 840
Db 37649 TCAAGGTTATCGCTTCTGCGGTGATCTGAAGATTAAATCGGCTTACACCGTCAAAAA 37590
QY 841 AACGGCGCTTTTATAGCGCGTATTTTATTTTCAACCTTATTTCCAGATACGTAATCATC 900
Db 37589 AACGGCGCTTTTATAGCGCGTATTTTATTTTCAACCTTATTTCCAGATACGTAATCATC 37530
QY 901 GTCCGTTGTAATCTTCTACTGCGCTT 927
Db 37529 GTCCGTTGTAATCTTCTACTGCGCTT 37503

RESULT 3
ECOW93/c
LOCUS Escherichia coli K-12 chromosomal region from 92.8 to 100.1 minutes.
DEFINITION Escherichia coli K-12 chromosomal region from 92.8 to 100.1 minutes.
ACCESSION U14003
VERSION U14003.1 GI:1263172
KEYWORDS Escherichia coli
SOURCE Escherichia coli
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 338534)
AUTHORS Burland, V., Plunkett, G. III, Sofia, H. J., Daniels, D. L. and Blattner, F. R.
TITLE Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.8 through 100 minutes
JOURNAL Nucleic Acids Res. 23 (12), 2105-2119 (1995)
PUBMED 7610040
REFERENCE 2 (bases 1 to 338534)
AUTHORS Plunkett, G. III.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-1994) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459
COMMENT On Apr 17, 1996 this sequence version replaced gi:536929.
This sequence was determined as part of the E. coli Genome Project (Frederick R. Blattner, director) at the University of Wisconsin-Madison. Supported by award HG00301 from the NIH Human Genome Project. The entire sequence was independently determined from E. coli MG1655; overlaps with other sequence determinations are annotated. The start of this entry overlaps the end of the entry ECOW89 (U00006) by 1885 bp.

Location/Qualifiers
1. 338534
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="K-12"
/sub_strain="MG1655"
/db_xref="taxon:562"
/map="92.8 to 100 minutes"
/notes="This sequence comprises the following lambda clones: DD928(EC27-278), DD930(EC30-188), DD933 (EC27-1320), DD935 (EC30MM32), DD937 (EC27-1070), DD941 (EC17-142), DD945 (EC17-8), DD947 (EC34A-34), DD949 (EC22-169), DD952 (EC27-297), DD953 (EC27-409), DD956 (EC17-103), DD958 (EC30MM1), DD960 (EC30MM2), DD962 (EC21-104), DD965 (EC23A-40), DD968 (EC30K660A-4pp), DD970 (EC19-202), DD974 (EC19-61), DD975 (EC18-233), DD977 (EC30MM7), DD980 (EC27-1151), DD981 (EC18-3), DD984 (EC17-101), DD987 (EC18-115), DD990 (EC17-136), DD992 (EC18-282), DD995 (EC17-115), DD997 (EC18-113); M13mp19 or Janus vectors were used for subcloning"
1. 1885
/note="1885 bp overlap with end of GenBank Accession Number U00006 (ECOW89)"
Complement (<1. 398)
/note="ORF f326 of GenBank Accession Number U00006 (ECOW89)"
/codon_start=1
/transl_table=11
/label_ORF_f132p
/protein_id="AAA96985.1"
/db_xref="GI:1263173"
/translation="MGFTTRVKESEBKKPFNFALFMDKYGTFFILAIIVAFGSLSP EYFTTNNTIOTFVQSVTVLIGMEFFAILVAGIDLSVGAIALSGMVTAKMLAGV DPFLAAMTGGVLVGGALGANGCLVNTGL"
277. 397
/note="predicted bend of 75.10 degrees"
complement (377. 1909)
/note="ORF f510"
/codon_start=1
/transl_table=11
/label_ORF_f510
/protein_id="AAA96986.1"
/db_xref="GI:536931"
/translation="MATPYISMAGIKSGFGVHALKSVNLTVYPGEIHALGENGAGK STLKMLSGIHEPTKTTINNINSLKDLHQAQGLIYQSLSDIETLVLENLY IGRHLTKICGVNIIIDWREVRVRAAMLLRVGLQVLDKLVANLSISHKOMLEIAKTL MLDKAVIMDEPTSSLTKNVDYFLIMNQLRKEGTALVISHKLAERRRICDRVTVM KGSVSCGIVSDNDDIVRLMVGRELQNFNKNVSNLAHETVFEVRVSRDR KKKRDI8FVCRGEILFAGLVGSGRTLMNCLFQVDRKAGEIRLNGKDISPRSLD AVKKGMAVITSERRDNGFFPNFSAQNMAISRSLKGGYKGMGLFHEVDQRTAENQ RELALAKHSVNQNTBLSGNQOKVLISKWLCCEPVIIFDEPTRGIDVGAKEIYK VMRQLADDGKVLVMSSELPIITVCDRIAVFCSEGRLTQILTRDDNSEEIMAWALP Q8"
809. 931
/note="predicted bend of 74.75 degrees"
1768. 1889
/note="predicted bend of 75.87 degrees"
complement (1983. 2010)
/note="promoter matrix score of 55; putative; resembles flh-regulated promoter"
complement (1985. 2021)
/standard_name="REP; repetitive extragenic palindromic element"
/note="contains 1 REP sequence"
complement (2036. 2971)
/note="ORF_f311"
/codon_start=1
/transl_table=11
/label_ORF_f311
/protein_id="AAA96987.1"
/db_xref="GI:536932"
/translation="MNKYLKYPFSGTLVGLMLSTAFAPAAAEYAVLVKTSNPFWDMMKK GIEDEAKTLGVSVDIFAPSPSEGDQSQQLFEDLSNKNYKGIAPAPLSPVNLVMPVAR


```
Db 330579 GAAATGACCGCGGTGAGCTGAAACCGCACGACCGTACTGTAGACGTGACGATCCGCG 330520
Qy 721 TATTCGTAAACATTTTCGAATCTACGCGCGGATACGCGGGAATCATCGCCACCATTCACGG 780
Db 330519 TATTCGTAAACATTTTCGAATCTACGCGCGGATACGCGGGAATCATCGCCACCATTCACGG 330460
Qy 781 TGAAGGTTATCGCTTCTGCGGTGATCTGGAAGATTAAATCGGCTTTTACCAACCGTCAAAAAA 840
Db 330459 TGAAGGTTATCGCTTCTGCGGTGATCTGGAAGATTAAATCGGCTTTTACCAACCGTCAAAAAA 330400
Qy 841 AACGGCGCTTTTAGCGCGTATTTTATTTTCAACTTATTTCCAGATACGTAACTCATC 900
Db 330399 AACGGCGCTTTTAGCGCGTATTTTATTTTCAACTTATTTCCAGATACGTAACTCATC 330340
Qy 901 GTCCGTTGTAACCTCTTTTACTGCGCTTT 927
Db 330339 GTCCGTTGTAACCTCTTTTACTGCGCTTT 330313
```

RESULT 4
AE005174 55/c
WPCOMMENT

Sequence split into 56 fragments LOCUS AE005174 Accession AE005174

Fragment Name	Begin	End
AE005174_00	1	110000
AE005174_01	100001	210000
AE005174_02	200001	310000
AE005174_03	300001	410000
AE005174_04	400001	510000
AE005174_05	500001	610000
AE005174_06	600001	710000
AE005174_07	700001	810000
AE005174_08	800001	910000
AE005174_09	900001	1010000
AE005174_10	1000001	1110000
AE005174_11	1100001	1210000
AE005174_12	1200001	1310000
AE005174_13	1300001	1410000
AE005174_14	1400001	1510000
AE005174_15	1500001	1610000
AE005174_16	1600001	1710000
AE005174_17	1700001	1810000
AE005174_18	1800001	1910000
AE005174_19	1900001	2010000
AE005174_20	2000001	2110000
AE005174_21	2100001	2210000
AE005174_22	2200001	2310000
AE005174_23	2300001	2410000
AE005174_24	2400001	2510000
AE005174_25	2500001	2610000
AE005174_26	2600001	2710000
AE005174_27	2700001	2810000
AE005174_28	2800001	2910000
AE005174_29	2900001	3010000
AE005174_30	3000001	3110000
AE005174_31	3100001	3210000
AE005174_32	3200001	3310000
AE005174_33	3300001	3410000
AE005174_34	3400001	3510000
AE005174_35	3500001	3610000
AE005174_36	3600001	3710000
AE005174_37	3700001	3810000
AE005174_38	3800001	3910000
AE005174_39	3900001	4010000
AE005174_40	4000001	4110000
AE005174_41	4100001	4210000
AE005174_42	4200001	4310000
AE005174_43	4300001	4410000
AE005174_44	4400001	4510000
AE005174_45	4500001	4610000
AE005174_46	4600001	4710000
AE005174_47	4700001	4810000

AE005174_48	4800001	4910000
AE005174_49	4900001	5010000
AE005174_50	5000001	5110000
AE005174_51	5100001	5210000
AE005174_52	5200001	5310000
AE005174_53	5300001	5410000
AE005174_54	5400001	5510000
AE005174_55	5500001	5528445
Continuation (56 of 56) of AE005174 from base 5500001 (AE005174 Escherichia coli O157:H7)		
Query Match 99.3%; Score 920.6; DB 1; Length 28445;		
Best Local Similarity 99.6%; Pred. No. 5.7e-260;		
Matches 923; Conservative 0; Mismatches 4; Indels 0; Gaps 0;		
Qy 1	GTCTATGTTACGCCCATCATGTTAATTTGCACATGTCATCAGGCAGGTTCAGGACCTTTTGT	60
Db 271199	GTCTATGTTACGCCCATCATGTTAATTTGCACATGTCATCAGGCAGGTTCAGGACCTTTTGT	271140
Qy 61	ACTTCCTGTTTCGATTTAGTTGGCAATTTAGTTAGCAAAATGCGAGACCCCGCACTTCT	120
Db 271139	ACTTCCTGTTTCGATTTAGTTGGCAATTTAGTTAGCAAAATGCGAGACCCCGCACTTCT	27080
Qy 121	TATCGTTGAACGACGAGTTGTTAAACGCAACACGTTGAAAGATATTTTCGAGCGGAAG	180
Db 27079	TATCGTTGAACGAGTTGTTAAACGCAACACGTTGAAAGATATTTTCGAGCGGAAG	27020
Qy 181	CTATGATGTTTTCGAAGCAGACAGATGGCGCGGAATGTCATCAGATCCTCTCTGAATATGA	240
Db 27019	CTATGATGTTTTCGAAGCAGACAGATGGCGCGGAATGTCATCAGATCCTCTCTGAATATGA	26960
Qy 241	CATCAACCTGTTGATCATGATATCAATCTCCCGGTAAAGAACGGTCTTCTGTTAGCGCG	300
Db 26959	CATCAACCTGTTGATCATGATATCAATCTCCCGGTAAAGAACGGTCTTCTGTTAGCGCG	26900
Qy 301	TGAACTGCGGAGCAGCGGAATGTTGCGTTGATGTTCTGCTGCGGTGCAACGAAGT	360
Db 26899	TGAACTGCGGAGCAGCGGAATGTTGCGTTGATGTTCTGCTGCGGTGCAACGAAGT	26840
Qy 361	CGATAAAATTTCTCGGCCCTCGAAATCGGTGTCAGATGACTACATCAACAAACCGTTCAACCC	420
Db 26839	CGATAAAATTTCTCGGCCCTCGAAATCGGTGTCAGATGACTACATCAACAAACCGTTCAACCC	26780
Qy 421	CGGTGAACCTGACGATTCGTGCAACCTACTGTCCCGTACCATGAAATCTGGGTACTGT	480
Db 26779	CGGTGAACCTGACGATTCGTGCAACCTACTGTCCCGTACCATGAAATCTGGGTACTGT	26720
Qy 481	CAGCGAAGACGTGCTAGCGTTGAAAGCTACAAGTTCAATGGTTGGGAACCTGGACATCAA	540
Db 26719	CAGCGAAGACGTGCTAGCGTTGAAAGCTACAAGTTCAATGGTTGGGAACCTGGACATCAA	26660
Qy 541	CAGCGTTTCGTTGATCGGCCCTCGATGGCGCAGCAGTACAAGCTGCGCGCAGCGAGTCCG	600
Db 26659	CAGCGTTTCGTTGATCGGCCCTCGATGGCGCAGCAGTACAAGCTGCGCGCAGCGAGTCCG	26600
Qy 601	CGCCATGTTTCACTTCTGTGAAAAACCCAGGCAAAATTCAGTCCCGTGTGAACTGCTGAA	660
Db 26599	CGCCATGTTTCACTTCTGTGAAAAACCCAGGCAAAATTCAGTCCCGTGTGAACTGCTGAA	26540
Qy 661	GAAATGACCGCGCGTGAAGCTGAAACCGCAACCGTACTGTAGACGTGACGATCCGCG	720
Db 26539	GAAATGACCGCGCGTGAAGCTGAAACCGCAACCGTACTGTAGACGTGACGATCCGCG	26480
Qy 721	TATTGTTAAACATTTTCGAATCTACGCCGGATACGCCGGAAATCATCGCCACCATTCACGG	780
Db 26479	TATTGTTAAACATTTTCGAATCTACGCCGGATACGCCGGAAATCATCGCCACCATTCACGG	26420
Qy 781	TGAAGGTTATCGCTTCTCGGTGATCTGGAAGATTAATCGGCTTTTACCGTCAAAAAA	840
Db 26419	TGAAGGTTATCGCTTCTCGGTGATCTGGAAGATTAATCGGCTTTTACCGTCAAAAAA	26360
Qy 841	AACGGCGCTTTTAGCGCGTATTTTATTTTCAACCTTATTTCCAGATACGTAACTCATC	900
Db 26359	AACGGCGCTTTTAGCGCGTATTTTATTTTCAACCTTATTTCCAGATACGTAACTCATC	26300

QY 901 GTCCGTTGTAACCTCTTTACTGGCTTT 927
DB 26299 GTCCGTTGTAACCTCTTTACTGGCTTT 26273

RESULT 5

BA000007_54/c

WPCOMMENT

Sequence split into 55 fragments LOCUS BA000007 Accession BA000007

Fragment Name Begin End

BA000007_00 1 110000

BA000007_01 100001 210000

BA000007_02 200001 310000

BA000007_03 300001 410000

BA000007_04 400001 510000

BA000007_05 500001 610000

BA000007_06 600001 710000

BA000007_07 700001 810000

BA000007_08 800001 910000

BA000007_09 900001 1010000

BA000007_10 1000001 1110000

BA000007_11 1100001 1210000

BA000007_12 1200001 1310000

BA000007_13 1300001 1410000

BA000007_14 1400001 1510000

BA000007_15 1500001 1610000

BA000007_16 1600001 1710000

BA000007_17 1700001 1810000

BA000007_18 1800001 1910000

BA000007_19 1900001 2010000

BA000007_20 2000001 2110000

BA000007_21 2100001 2210000

BA000007_22 2200001 2310000

BA000007_23 2300001 2410000

BA000007_24 2400001 2510000

BA000007_25 2500001 2610000

BA000007_26 2600001 2710000

BA000007_27 2700001 2810000

BA000007_28 2800001 2910000

BA000007_29 2900001 3010000

BA000007_30 3000001 3110000

BA000007_31 3100001 3210000

BA000007_32 3200001 3310000

BA000007_33 3300001 3410000

BA000007_34 3400001 3510000

BA000007_35 3500001 3610000

BA000007_36 3600001 3710000

BA000007_37 3700001 3810000

BA000007_38 3800001 3910000

BA000007_39 3900001 4010000

BA000007_40 4000001 4110000

BA000007_41 4100001 4210000

BA000007_42 4200001 4310000

BA000007_43 4300001 4410000

BA000007_44 4400001 4510000

BA000007_45 4500001 4610000

BA000007_46 4600001 4710000

BA000007_47 4700001 4810000

BA000007_48 4800001 4910000

BA000007_49 4900001 5010000

BA000007_50 5000001 5110000

BA000007_51 5100001 5210000

BA000007_52 5200001 5310000

BA000007_53 5300001 5410000

BA000007_54 5400001 5498450

Continuation (55 of 55) of BA000007 from base 5400001 (BA000007 Escherichia coli O157:H7)

Query Match 99.3%; Score 920.6; DB 1; Length 98450;
Best Local Similarity 99.6%; Pred. No. 7.3e-260;
Matches 923; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GTCATGTTACGCCGATCATGTTAAATTTGCAGCATGATCAGGCAGGTCAGGACTTTTGT 60

DB 97204 GTCATGTTACGCCGATCATGTTAAATTTGCAGCATGATCAGGCAGGTCAGGACTTTTGT 97145
QY 61 ACTTCCTGTTTCGATTTAGTTGGCAATTTAGGTAGCAAAACATGACAGCCCGCACATTTCT 120
DB 97144 ACTTCCTGTTTCGATTTAGTTGGCAATTTAGGTAGCAAAACATGACAGCCCGCACATTTCT 97085
QY 121 TATCGTTGAAGACGAGTTGGTAACACGCAACCGTTGAAAAGTATTTTCGAAGCGGAAGG 180
DB 97084 TATCGTTGAAGACGAGTTGGTAACACGCAACCGTTGAAAAGTATTTTCGAAGCGGAAGG 97025
QY 181 CTATGATGTTTTCGAAGCGACAGATGGCGCGGAATGTCATCAGATTCCTCTCTGAATATGA 240
DB 97024 CTATGATGTTTTCGAAGCGACAGATGGCGCGGAATGTCATCAGATTCCTCTCTGAATATGA 96965
QY 241 CATCAACCTGATGATCATGATATCAATCTGCCGGGTAAAGAACGGTCTTCTGTTAGCCGG 300
DB 96964 CATCAACCTGATGATCATGATATCAATCTGCCGGGTAAAGAACGGTCTTCTGTTAGCCGG 96905
QY 301 TGAACCTGCGAGCAGCGCAATGTTGCTGATGTTCTGACTGCGCGGTGACACGAAAT 360
DB 96904 TGAACCTGCGAGCAGCGCAATGTTGCTGATGTTCTGACTGCGCGGTGACACGAAAT 96845
QY 361 CGATAAAATCTCGGCCCTCGAAATCGGTGCGAGATGACTACATCAACCAACCGTTCAACCC 420
DB 96844 CGATAAAATCTCGGCCCTCGAAATCGGTGCGAGATGACTACATCAACCAACCGTTCAACCC 96785
QY 421 GCGTGAACCTGACGATTCGTGACGCAACCTACTGTCGGTACCATGATCTCGGTACTGTT 480
DB 96784 GCGTGAACCTGACGATTCGTGACGCAACCTGTCCTGCGTACCATGATCTCGGTACTGTT 96725
QY 481 CAGCGAAGACGTCGTAGCGTTGAAAGCTACAAGTTCAATGTTGGGAACCTGGACATCAA 540
DB 96724 CAGCGAAGACGTCGTAGCGTTGAAAGCTACAAGTTCAATGTTGGGAACCTGGACATCAA 96665
QY 541 CAGCGGTTTCGTTGATCGGCCCTGATGCGGAGCAGTACAAGTGGCGCGCAGCGAGTTCCG 600
DB 96664 CAGCGGTTTCGTTGATCGGCCCTGATGCGGAGCAGTACAAGTGGCGCGCAGCGAGTTCCG 96605
QY 601 CGCCATGCTTCACTTCTGTGAAAACCCAGGCAAAATTCAGTCCCGTGTGACTGCTGAA 660
DB 96604 CGCCATGCTTCACTTCTGTGAAAACCCAGGCAAAATTCAGTCCCGTGTGACTGCTGAA 96545
QY 661 GAAATGACCGCGCGTGTAGCTGAAACCGCAGCAGCGTACTGTAGACGTGACGATCCGCGG 720
DB 96544 GAAATGACCGCGCGTGTAGCTGAAACCGCAGCAGCGTACTGTAGACGTGACGATCCGCGG 96485
QY 721 TATTCGTAACATTTTCGAATCTACCGCGGATACCGCGGAATCATTCGCCACATTCACGG 780
DB 96484 TATTCGTAACATTTTCGAATCTACCGCGGATACCGCGGAATCATTCGCCACATTCACGG 96425
QY 781 TGAAGGTTATCGCTTCTGCGGTGATCTGGAAGATTATCGGCTTTACACCGTCAAAAA 840
DB 96424 TGAAGGTTATCGCTTCTGCGGTGATCTGGAAGATTATCGGCTTTACACCGTCAAAAA 96365
QY 841 AACCGCGCTTTTACGCGCGTTTTTATTTTCAACCTTATTTCCAGATACGTAATCATC 900
DB 96364 AACCGCGCTTTTACGCGCGTTTTTATTTTCAACCTTATTTCCAGATACGTAATCATC 96305
QY 901 GTCCGTTGTAACCTTCTTTTACTGGCTTT 927
DB 96304 GTCCGTTGTAACCTTCTTTTACTGGCTTT 96278

RESULT 6

AE016772/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AE016772 Escherichia coli CFT073 section 18 of 18 of the complete genome.
Escherichia coli CFT073
Escherichia coli CFT073
Escherichia coli CFT073
linear BCT 24-DEC-2002

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.

REFERENCE AUTHORS

1 (bases 1 to 86898)
Welch,R.A., Burland,V., Plunkett,G.D. III, Redford,P., Roesch,P., Rasko,D.A., Buckles,E.L., Liou,S.-R., Boutin,A., Hackett,J., Stroud,D., Mayhew,G.P., Rose,D.J., Zhou,S., Schwartz,D.C., Perna,N.T., Mobley,H.L.T., Donnenberg,M.S. and Blattner,F.R.
Extensive Mosaic Structure Revealed by the Complete Genome Sequence of Uropathogenic Escherichia coli
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 17020-17024 (2002)
12471157

TITLE

JOURNAL PUBMED

REFERENCE AUTHORS

2 (bases 1 to 86898)
Welch,R.A., Burland,V., Plunkett,G.D. III, Redford,P., Roesch,P., Rasko,D.A., Buckles,E.L., Liou,S.-R., Boutin,A., Hackett,J., Stroud,D., Mayhew,G.P., Rose,D.J., Zhou,S., Schwartz,D.C., Perna,N.T., Mobley,H.L.T., Donnenberg,M.S. and Blattner,F.R.
Direct Submission
Submitted (20-JUN-2002) Genetics Laboratory, University of Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
Location/Qualifiers
1. .86898
/organism="Escherichia coli CFT073"
/mol_type="genomic DNA"
/strain="CFT073"
/db_xref="taxon:199310"
complement(152..1495)
/gene="gntP"
/locus_tag="c5401"
complement(152..1495)
/gene="gntP"
/locus_tag="c5401"
/function="transport; Transport of small molecules: Carbohydrates, organic acids, alcohols"
/note="Glucanate permease 3; Escherichia coli K-12 ortholog: b4321; Escherichia coli O157:H7 ortholog: z5919"
/codon_start=1
/transl_table=1
/product="High-affinity gluconate transporter"
/protein_id="AAN83823.1"
/db_xref="GI:26111642"

FEATURES source

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

/gene="uxuB"
/locus_tag="c5403"
3100..4560
/gene="uxuB"
/locus_tag="c5403"
/EC_number="1.1.1.57"
/function="enzyme; Degradation of small molecules: Carbon compounds"
/note="Escherichia coli K-12 ortholog: b4323; Escherichia coli O157:H7 ortholog: z5921"
/codon_start=1
/transl_table=11
/product="D-mannanate oxidoreductase"
/protein_id="AAN83825.1"
/db_xref="GI:26111644"
/translation="MTTIVDSNLVAPSPWDSRLESRIVHLGGCAFRAHQALYTHH
LLESTDSWGICEVNLPGNDRLVLIENLKQQLLYTVBAKGAESTELKIGSMKALH
PEIDCGEILNMAPQTAIVSLVTERGYCAADAGSDLDNNPLIKHDLNPTAPK
AIGYIVELRLRREKKAFTVSCDNRVRENGHVAKAVLGLAQARDPQLAAITENV
TFPCTVMDRIVPAATPETLQBIADQLGVYDCAICBPFQWVIENFVNGRPMWDKV
GAQFVADVPPPEMKMLRMLGSHFLAYLGLVGYETIADTWTNPAYKAAFPALMQE
OAPTLSMEGTDLNAYATLLIERPSLSLHRTWQIAMDGSKLPQRLDPVRLHLQ
GGSWRHLALGVAGWMRYTQGVDEQNALDVVDPMMAEFQKINAQTGGADRVKALLGLS
GIFADDLPQNADFVGAVTAAVYQLCERGARECVAAI"
4775..5548
/gene="uxuR"
/locus_tag="c5404"
4775..5548
/gene="uxuR"
/locus_tag="c5404"
/function="regulator; Degradation of small molecules: Carbon compounds"
/note="Escherichia coli K-12 ortholog: b4324; Escherichia coli O157:H7 ortholog: z5922"
/codon_start=1
/transl_table=11
/product="Uxu operon transcriptional regulator"
/protein_id="AAN83826.1"
/db_xref="GI:26111645"
/translation="MKSATSAORPQVQVGMIRDLIIKTPNPGERLPPEREIEMLD
VTRTVREALIMLEIKGLVEVRGAGIYVLDSSGHADSDANVCNDAGPPELLOAR
QLLESNTAFALQATREDIVMKAQALQERELASAPSGESGDMQFHLAIABATH
NSMLVELFRSQWRENNPMWILQSHLDLSLYKRWLGDHKQILAAALIKKDAARAKL
AMQHLNKKQRLLEFNSNVDIYFDGVLFDSPWLDKVDAA"
5953..6345
/gene="yjid"
/locus_tag="c5405"
5953..6345
/gene="yjid"
/locus_tag="c5405"
/note="Escherichia coli K-12 ortholog: b4326; Escherichia coli O157:H7 ortholog: z5925"
/codon_start=1
/transl_table=11
/product="Hypothetical protein yjid"
/protein_id="AAN83827.1"
/db_xref="GI:26111646"
/translation="MRKMRQSVQTVLPESGNTLSLRDSVCRDLFQSFSPHPLP
ILLVSGPWEQGNQSKLQSWCYRQLRSALLFHEPRIAALQVNLKRAYCHELAISL
EMMLYHDDLEPLTFDLVWQKSGSWHRTMPQ"
complement(6393..6734)
/locus_tag="c5406"
complement(6393..6734)
/locus_tag="c5406"
/note="Residues 1 to 113 of 113 are 95.57 pct identical to residues 424 to 536 of 536 from GenPept.129 : >gb|AA161895.1| (AF289032) putative phosphoenolpyruvate-protein phosphoryltransferase Ppte [Escherichia coli]"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="AAN83828.1"

FEATURES		Location/Qualifiers	
source	1..1468	/organism="Escherichia coli"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:562"	
gene	98..814	/gene="dye"	
CDS	98..814	/gene="dye"	
		/codon_start=1	
		/transl_table=11	
		/protein_id="AAA23718.1"	
		/db_xref="GI:145818"	
EYDINLVMDINLPKNGILLAREQANVALMPLTRDNEVDKILGLEIGADDDYIT KPNPRELTIARNLSTNMLGTVSEERRSVESKFNGLWLDINSRLIGPDGEQYK LPRESEPRAMLHPCENPGKIQSRAELLKQWGRELKPDRYTDVYIRIRKRFESTPTD PEIATIHGEGRFCGDUEH			
ORIGIN	699 bp upstream of Acci site.		
Query Match	98.4%; Score 912; DB 1; Length 1468;		
Best Local Similarity	99.9%; Pred. No. 1.1e-257;		
Matches 923; Conservative	0; Mismatches 0; Indels 1; Gaps 1;		
QY	5	TGTTACGCCGATCATGTAAATTGCGATGTCATCAGGCAGGTCAAGGA-CTTTGTGACT	63
Db	1	TGTTACGCCGATCATGTAAATTGCGATGTCATCAGGCAGGTCAAGGCACCTTTGTGACT	60
QY	64	TCCTGTTTCGATTAGTTGGCAATTAGGTAGCAACATGCGAGCCCGGCACATCTTTAT	123
Db	61	TCCTGTTTCGATTAGTTGGCAATTAGGTAGCAACATGCGAGCCCGGCACATCTTTAT	120
QY	124	CCTTCAAGCAGTGTGTAACACGCAACAGCTTGAAGATATTTTGAAGCGGAGGCTA	183
Db	121	CCTTCAAGCAGTGTGTAACACGCAACAGCTTGAAGATATTTTGAAGCGGAGGCTA	180
QY	184	TGATGTTTTTGAAGCGACAGATGGCGGGAATGTCATCAGATCTCTCTGAATATGACAT	243
Db	181	TGATGTTTTTGAAGCGACAGATGGCGGGAATGTCATCAGATCTCTCTGAATATGACAT	240
QY	244	CAACTGTGTGATCATGGATATCAATCTCGCGGTAAGAACGGTCTTCTGTAGCGCGTGA	303
Db	241	CAACTGTGTGATCATGGATATCAATCTCGCGGTAAGAACGGTCTTCTGTAGCGCGTGA	300
QY	304	ACTGCGCAGCAGGGAATGTTGCGTGTCTGACTGCGCTGACACGAGTGA	363
Db	301	ACTGCGCAGCAGGGAATGTTGCGTGTCTGACTGCGCTGACACGAGTGA	360
QY	364	TAAATTTCTCGSCCTCGAAATCGGTGCGAGATGACTACATCACCAACCGTTCAACCCGCG	423
Db	361	TAAATTTCTCGSCCTCGAAATCGGTGCGAGATGACTACATCACCAACCGTTCAACCCGCG	420
QY	424	TGAATCTGACGATTCTGTGACGCAACTACTGTCCGTTACATGATCTGGGTACTGTGAC	483
Db	421	TGAATCTGACGATTCTGTGACGCAACTACTGTCCGTTACATGATCTGGGTACTGTGAC	480
QY	484	CGAAGACGTCTAGCTGTGAAGCTACAAGTTCAATGTTGGGAACCTGGACATCAACAG	543
Db	481	CGAAGACGTCTAGCTGTGAAGCTACAAGTTCAATGTTGGGAACCTGGACATCAACAG	540
QY	544	CGGTTCTGTTGATCGGCCCTTGATGGCGAGCAGTACAAAGCTGCGCGCAGCGAGTTCCGCGC	603
Db	541	CGGTTCTGTTGATCGGCCCTTGATGGCGAGCAGTACAAAGCTGCGCGCAGCGAGTTCCGCGC	600
QY	604	CATGCTTCACTTCTGTGAAAACCCAGGCAAAATTCAGTCCCGTGTGTAACCTGCTGAAGAA	663
Db	601	CATGCTTCACTTCTGTGAAAACCCAGGCAAAATTCAGTCCCGTGTGTAACCTGCTGAAGAA	660
QY	664	AATGACCGCGCTGAGCTGAACCGCAGCGTACTGTAGCGTGAAGATCCGCGCTAT	723
Db	661	AATGACCGCGCTGAGCTGAACCGCAGCGTACTGTAGCGTGAAGATCCGCGCTAT	720
QY	724	TCGTAACATTTTTCGAATCTAAGCCCGGATACGCCGGAATCATCGCCACCATTCACGGTGA	783

Db	721	TCGTAACATTTTCGAATCTACGCCGATACGCCGGAATCATCGCCACCATTCACGGTGA	78
QY	784	AGGTTATCGCTTCTCGCGGTGATCTGGAAGATTAATCGGCTTTTACCAACCGTCAAAAAAAC	843
Db	781	AGGTTATCGCTTCTCGCGGTGATCTGGAAGATTAATCGGCTTTTACCAACCGTCAAAAAAAC	840
QY	844	GGCGCTTTTATAGCGCGTTTTATTTTCAACCTTATTTCCAGATACGTAACCTCATCGTC	903
Db	841	GGCGCTTTTATAGCGCGTTTTATTTTCAACCTTATTTCCAGATACGTAACCTCATCGTC	900
QY	904	CGTTGTAACTTCTTTACTGGCTTT	927
Db	901	CGTTGTAACTTCTTTACTGGCTTT	924
RESULT 8			
AE005674_45/c			
WPCOMMENT			
Sequence split into 46 fragments			
	Fragment Name	Begin	End
	AE005674_00	1	110000
	AE005674_01	100001	210000
	AE005674_02	200001	310000
	AE005674_03	300001	410000
	AE005674_04	400001	510000
	AE005674_05	500001	610000
	AE005674_06	600001	710000
	AE005674_07	700001	810000
	AE005674_08	800001	910000
	AE005674_09	900001	1010000
	AE005674_10	1000001	1110000
	AE005674_11	1100001	1210000
	AE005674_12	1200001	1310000
	AE005674_13	1300001	1410000
	AE005674_14	1400001	1510000
	AE005674_15	1500001	1610000
	AE005674_16	1600001	1710000
	AE005674_17	1700001	1810000
	AE005674_18	1800001	1910000
	AE005674_19	1900001	2010000
	AE005674_20	2000001	2110000
	AE005674_21	2100001	2210000
	AE005674_22	2200001	2310000
	AE005674_23	2300001	2410000
	AE005674_24	2400001	2510000
	AE005674_25	2500001	2610000
	AE005674_26	2600001	2710000
	AE005674_27	2700001	2810000
	AE005674_28	2800001	2910000
	AE005674_29	2900001	3010000
	AE005674_30	3000001	3110000
	AE005674_31	3100001	3210000
	AE005674_32	3200001	3310000
	AE005674_33	3300001	3410000
	AE005674_34	3400001	3510000
	AE005674_35	3500001	3610000
	AE005674_36	3600001	3710000
	AE005674_37	3700001	3810000
	AE005674_38	3800001	3910000
	AE005674_39	3900001	4010000
	AE005674_40	4000001	4110000
	AE005674_41	4100001	4210000
	AE005674_42	4200001	4310000
	AE005674_43	4300001	4410000
	AE005674_44	4400001	4510000
	AE005674_45	4500001	4607203
Continuation (46 of 46) of AE005674 from base 4500001 (AE005674 Shigella flex			
Query Match 98.2%; Score 910.2; DB 1; Length 107203;			
Best Local Similarity 99.6%; Pred. No. 8.7e-257;			
Matches 923; Conservative 0; Mismatches 1; Gaps 1;			

Db	721	TCGTAACATTTTCGAATCTACGCCGATACGCCGGAATCATCGCCACCATTCACGGTGA	780
QY	784	AGGTTATCGCTTCTGCGGTGATCTGGAAGATTAATCGCCTTTACACCGTCAAAAAAAC	843
Db	781	AGGTTATCGCTTCTGCGGTGATCTGGAAGATTAATCGCCTTTACACCGTCAAAAAAAC	840
QY	844	GGCGCTTTTAGCGCCGTTTTTATTTTCAACCTTATTTCCAGATACCTAACTCATCGTC	903
Db	841	GGCGCTTTTAGCGCCGTTTTTATTTTCAACCTTATTTTCCAGATACCTAACTCATCGTC	900
QY	904	CGTTGTAACTTCTTTACTGGCTTT	927
Db	901	CGTTGTAACTTCTTTACTGGCTTT	924
RESULT 8			
AE005674 45/c			
WPCOMMENT			
Sequence split into 46 fragments LOCUS AE005674 Accession AE005674			
Fragment Name Begin End			
AE005674_00 1 110000			
AE005674_01 100001 210000			
AE005674_02 200001 310000			
AE005674_03 300001 410000			
AE005674_04 400001 510000			
AE005674_05 500001 610000			
AE005674_06 600001 710000			
AE005674_07 700001 810000			
AE005674_08 800001 910000			
AE005674_09 900001 1010000			
AE005674_10 1000001 1110000			
AE005674_11 1100001 1210000			
AE005674_12 1200001 1310000			
AE005674_13 1300001 1410000			
AE005674_14 1400001 1510000			
AE005674_15 1500001 1610000			
AE005674_16 1600001 1710000			
AE005674_17 1700001 1810000			
AE005674_18 1800001 1910000			
AE005674_19 1900001 2010000			
AE005674_20 2000001 2110000			
AE005674_21 2100001 2210000			
AE005674_22 2200001 2310000			
AE005674_23 2300001 2410000			
AE005674_24 2400001 2510000			
AE005674_25 2500001 2610000			
AE005674_26 2600001 2710000			
AE005674_27 2700001 2810000			
AE005674_28 2800001 2910000			
AE005674_29 2900001 3010000			
AE005674_30 3000001 3110000			
AE005674_31 3100001 3210000			
AE005674_32 3200001 3310000			
AE005674_33 3300001 3410000			
AE005674_34 3400001 3510000			
AE005674_35 3500001 3610000			
AE005674_36 3600001 3710000			
AE005674_37 3700001 3810000			
AE005674_38 3800001 3910000			
AE005674_39 3900001 4010000			
AE005674_40 4000001 4110000			
AE005674_41 4100001 4210000			
AE005674_42 4200001 4310000			
AE005674_43 4300001 4410000			
AE005674_44 4400001 4510000			
AE005674_45 4500001 4607203			
Continuation (46 of 46) of AE005674 from base 4500001 (AE005674 Shigella flexneri 2a str			
Query Match 98.2%; Score 910.2; DB 1; Length 107203;			
Best Local Similarity 99.6%; Pred. No. 8.7e-257;			
Matches 923; Conservative 0; Mismatches 3; Indels 1; Gaps 1;			


```

to residues 60 to 241 of 241 from Escherichia coli K-12 :
B4311"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP19417.1"
/db_xref="GI:30043698"
/translation="MESNTWNTIHDNKKENALNDVQEVVNYAIKLDQWTVRPQGLT
HFSNGTRYGPVKLSWDATKDLNFGIRYVDWKAYROQDLSGDMRDNRVHWGQYVT
YHNSDFTYPAWOTLLYSQNDYRYANHKWATENAFVLQYHMTPTDITPYEYDYLDRQ
GVYNGRDLNLSNSYRIGVSFKL"
4029. .5243
/genes="yjht"
/locus_tag="S4469"
4029. .5243
/genes="yjht"
/locus_tag="S4469"
/notes="residues 1 to 404 of 404 are 98.76 pct identical to
residues 1 to 404 of 404 from Escherichia coli K-12 :
B4310"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP19418.1"
/db_xref="GI:30043699"
/translation="MTTLTARVFTTAEIIRKTVIALVCHLNCRSQETVTWNKTIIMAL
AIMMASFAANASVLPETVPVFKSGAIDNDTVYIGLSAGTAWYKLDTOAKDKKWTFA
LAAPGGPREQATSAFIDGNLVFGIGIKNSEGLTQVFNVDVHKYNPKNYSVVKLMSHA
PMGMAGHTVFVHNGKAVYTGQVQNIENGYPEDLNEACKSDTAIDKINAHYFDKKAED
YFFNFKLLSPDSTQNSYAGSPWGTAGAAVNVKNGDKTWLNGEARPLGRLTDAVE
LDFGNLKNWKLDDVSSPDGVAGGFAGISNDLSLIFAGGAGFKGSRNYQNGKNAEHE
GLKKSYSSTDHLHWHNGKWDKSGELSQGRAYGVSLPWNNSLLIIGGETAGKRAVTDVSL
ISVKNKVTQVQ"
5530. .6288
/genes="yjh"
/locus_tag="S4470"
5530. .6288
/genes="yjh"
/locus_tag="S4470"
/notes="residues 1 to 252 of 252 are 99.60 pct identical to
residues 75 to 326 of 326 from Escherichia coli K-12 :
B4309"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP19419.1"
/db_xref="GI:30043700"
/translation="WQGHHPHPLATHQYQYTVGQALHIARKLLPFIIPDNAGILIVPC
CRGSAFTAGSEGTYSERHGASHDACRWGTDTPLYODLVSRTAALAKNPQNKFLGVC
WMOGFDLWTSYASHPOHFHNMVEAFPRDLKQHSQNLNITDAPWFCGDTTWYWKEN
PPhSYEAITYGNQNNVLANIIFVDFQOQOERGLTNAPDEDDPDLSTGYGYSRSPEN
WTTALRSSHFSFAARRGIISDRFVEAILQFWRER"
6742. .7509
/notes="insertion element"
/insertion_seq="IS1X1"
/complement(6757. .7260)
/locus_tag="S4471"
/complement(6757. .7260)
/locus_tag="S4471"
/function="IS, phage, Tn; Transposon-related functions"
/notes="residues 1 to 167 of 167 are 89.82 pct identical to
residues 1 to 167 of 167 from GenPept :
>gb|AAD44742.1|AF141323_13 (AF141323) Tnph (Shigella
flexneri)"
/codon_start=1
/transl_table=11
/product="IS1 orfB"
/protein_id="AAP19420.1"
/db_xref="GI:30043701"
/translation="MSROCTHYGRWPHQHGFTSLKKLRPQSVTSRIQPSDVLVCAEMD
EHWGVGAKSRQRLVYAYDRIRTVVAHVFGERTLATLERLLLSLAFEVVWMTDG
WPLYESCLGKGLKHLVTSKRYTQRIERNLNLKQHLARLVKRSLSFSKVELHDKVIGHY
LNKHYQ"

```

```

gene
CDS
complement(7179. .7454)
/locus_tag="S4472"
complement(7179. .7454)
/locus_tag="S4472"
/function="IS, phage, Tn; Transposon-related functions"
/notes="residues 1 to 91 of 91 are 98.90 pct identical to
residues 1 to 91 of 91 from GenPept :
>gb|AAK18550.1|AF348706_239 (AF348706) iso-IS1 orfA
[Shigella flexneri]"
/codon_start=1
/transl_table=11
/product="IS1 orfA"
/protein_id="AAP19421.1"
/db_xref="GI:30043702"
/translation="WASISIRPCSCSAITEGVVRNGKSTAGHLRYLCSHCRKTKWQLQFT
YTASQPGTHQKIIDMANMNVGCRASARIMGVGLNVLRLHKNSGRSR"
7549. .8244
/locus_tag="S4473"
7549. .8244
/locus_tag="S4473"
/function="regulatory"
/notes="residues 1 to 231 of 231 are 81.46 pct identical to
residues 30 to 261 of 261 from GenPept : >gb|AAI22652.1|
(AE008876) putative regulatory protein, deor family
[Salmonella typhimurium LT2]"
/codon_start=1
/transl_table=11
/product="putative deor family transcriptional regulator"
/protein_id="AAP19422.1"
/db_xref="GI:30043703"
/translation="MLEVSVMTIRRDHLHQEDPLPLTLGGYIVMVNKPAPSMPIVHD
VPKRRDLPITAILAAGVNVENDLIIFDNGQEIPLVIMIDAITFTGICYSHRVFA
LNEKPNVTAILCGGYRARSDAFYDASNSPLDSLNPRIKIFISAGVNVHFGVSNFP
EDLATKRVNMRGLRKILLARHALFDEVASLASLAPISAFDVLISDRPLPADVYTHCQN
GSVKLIIPDSEDE"
complement(8345. .8461)
/locus_tag="S4474"
complement(8345. .8461)
/locus_tag="S4474"
/notes="residues 4 to 38 of 38 are 88.57 pct identical to
residues 29 to 63 of 96 from GenPept : >emb|CAC50649.1|
(AX188968) unnamed protein product [Escherichia coli]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP19423.1"
/db_xref="GI:30043704"
/translation="MWAFRRALNVEDILCLPIQDCDKSHIWLVLVKDQRLR"
Query Match 98.2%; Score 910.2; DB 1; Length 225944;
Best Local Similarity 99.6%; Pred. No. 1e-256;
Matches 923; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 GTCATGTTACGCCGATCATGTTAAATTCGACATGTCATCAGGACAGGACATTTGT 60
Db 224698 GTCATGTTACGCCGATCATGTTAAATTCGACATGTCATCAGGACAGGACATTTGT 224639
QY 61 ACTTCCTGTTTCGATTTAGTTGGCAATTTAGTAGCAACAATGACAGACCCGACATCT 120
Db 224638 ACTTCCTGTTTCGATTTAGTTGGCAATTTAGTAGCAACAATGACAGACCCGACATCT 224579
QY 121 TATCGTTCAAGACAGGTTGGTAAACACGCAACACCGTGTGAAAAGTATTTTCGAAGCGGAAGG 180
Db 224578 TATCGTTCAAGACAGGTTGGTAAACACGCAACACCGTGTGAAAAGTATTTTCGAAGCGGAAGG 224519
QY 181 CTATGATGTTTTCGAAGCGACAGATGCGCGGGAATATGATCATGATCTCTCTGAATATGA 240
Db 224518 CTATGATGTTTTCGAAGCGACAGATGCGCGGGAATATGATCATGATCTCTCTGAATATGA 224459
QY 241 CATCAACTGGTGATCATGGATATCAATCTGCCGGGTGAAGACGGTCTTCTGTTAGCGCG 300
Db 224458 CATCAACTGGTGATCATGGATATCAATCTGCCGGGTGAAGACGGTCTTCTGTTAGCGCG 224399

```

JOURNAL	Submitted (25-OCT-2001) Submitted on behalf of the Salmonella sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK E-mail: parkhill@sanger.ac.uk
COMMENT	Notes: Details of S. typhi sequencing at the Sanger Centre are available on the World Wide Web. (URL, http://www.sanger.ac.uk/Projects/s_typhi/).
FEATURES	Location/Qualifiers
Source	1. 88037 /organism="Salmonella enterica subsp. enterica serovar Typhi" /mol_type="genomic DNA" /strain="CT18" /db_xref="taxon:90370" 482_..1090 /gene="STY4858" 482_..1090 /gene="STY4858" /note="Similar to Streptomyces coelicolor hypothetical protein SC7H1.19 TR:O54178 (EMBL:AL021411) (182 aa) fasta scores: E(): 1.4e-10, 29.6% id in 169 aa" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="CAD03347.1" /db_xref="GI:16505619" /db_xref="InterPro:IPR009412" /db_xref="UniProt/TREMBL:Q8Z0Y5" /translation="MKVITVTPGVYRIAKRCPSCSVKRDFTPSGAFRVNSOKYLDVMSYKTCFHCIDTWNISFLSRPLVSKIRNLYCRLMANDAATVOYFAYDNAILKRNNAELSGQDPFHIOERWLVSITASHKQSVSRISFPQVSLLSILKKQLLSAAEIKRRIE TGOISGVYTMKMLSKRLKNAKYDLQLSVETLYDRRRIVLTRLR" 1451_..2362 /gene="STY4859" 1451_..2362 /gene="STY4859" /note="N-terminus is similar to Salmonella typhimurium IICA ilca TR:Q9RPB6 (EMBL:AF164435) (74 aa) fasta scores: E(): 3e-27, 98.6% id in 74 aa, indicated by promoterless reporter gene insertion to be induced upon bacterial invasion of phagocytic and nonphagocytic cells (but having no effect on virulence). The adjacent 70 aa is highly similar to the neighbouring CDS Salmonella typhimurium hypothetical protein TR:Q9RPB5 (EMBL:AF164435) (70 aa) fasta scores: E(): 1.9e-20, 91.4% id in 70 aa" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="CAD03348.1" /db_xref="GI:16505620" /db_xref="UniProt/TREMBL:Q8Z0Y4" /translation="MHTFRPYLLRHSLLVLEDPLEIREQIILLINILNCSSFFYDM TLYCVNHSDEVYRRICKTLRKKEYLFTYAHSTSYLDEMSNLLKTKDKRKHDTIT ELAFNYIDTYLRYEVTGLPEPKAISLANIIPHEISLKYRYENGRVLRRLIKRLN ICFYLYSPGYGFVEYDLMENRMLNDPACVRECHAVFRSVLLRIHRSKSTIVYHE ODSLNTLNLMLWARGVISAQVYVHFHFLIDVLESEIFLPWAPKSHHHHAWMLRISE LACSYYLTERTSIFFLQRAEDGVAP" 2573_..3034 /gene="STY4860" 2573_..3034 /gene="STY4860" /note="Similar to Neisseria meningitidis putative periplasmic protein NMA0993 TR:CA884263 (EMBL:AL162754) (159 aa) fasta scores: E(): 1.2e-12, 34.4% id in 160 aa and Neisseria meningitidis hypothetical protein NN00783 TR:AAF41196 (EMBL:AE002432) (159 aa) fasta scores: E(): 1.5e-12, 34.4% id in 160 aa. Also similar to Haemophilus ducreyi hypothetical protein TR:AAF33778 (EMBL:AF219260) (149 aa) fasta scores: E(): 9.2e-11, 31.5% id in 146 aa and Treponema pallidum hypothetical protein TP0895 TR:O83865 (EMBL:AE001259) (168 aa) fasta scores: E(): 6.9e-09, 33.6% id in 134 aa. Contains a probable
301	TGAACCTGCGGACGAGCGGAATGTTGCGTTGATGTTCTGTACTGCGCGTGACACGAAGT 360
224398	TGAACCTGCGGACGAGCGGAATGTTGCGTTGATGTTCTGTACTGCGCGTGACACGAAGT 224339
361	CGATAAAATCTCGGCTCGAAATCGGTGCGAGATGACTACATCACCACCGTTCAACCC 420
224338	CGATAAAATCTCGGCTCGAAATCGGTGCGAGATGACTACATCACCACCGTTCAACCC 224279
421	CGGTGAACCTGACGATTCGTGACGCAACCTACTGTCGCGTACCACGATGAATCTGGGTACTGT 480
224278	CGGTGAACCTGACGATTCGTGACGCAACCTGTCGCGTACCACGATGAATCTGGGTACTGT 224219
481	CAGCGAAGACGCTCGTAGCGTTTGAAGCTCAACAAGTTCAATGGTTGGGAATCGGACATCAA 540
224218	CAGCGAAGACGCTCGTAGCGTTTGAAGCTCAACAAGTTCAATGGTTGGGAATCGGACATCAA 224159
541	CAGCGCTTGCTGATCGGCGCTTGATGGCGAGCAGTACAAGCTGCCGCGCAGCGAGTTCCG 600
224158	CAGCGCTTGCTGATCGGCGCTTGATGGCGAGCAGTACAAGCTGCCGCGCAGCGAGTTCCG 224099
601	CGCCATGCTTCACTTCTGTGAAAACCCAGGCAAAATTCAGTCCGCTGCTGAACCTGCTGAA 660
224098	CGCCATGCTTCACTTCTGTGAAAACCCAGGCAAAATTCAGTCCGCTGCTGAACCTGCTGAA 224039
661	GAATAATGACCGCGCTGAGCTGAAACCGCACGACGCTACTGTAGACGTGACGATCCGCCG 720
224038	GAATAATGACCGCGCTGAGCTGAAACCGCACGACGCTACTGTAGACGTGACGATCCGCCG 223979
721	TATTCGTAAACATTTTCAATCTACGCGCGGATACGCGGAAATCATCTGCCACCATTCACGG 780
223978	TATTCGTAAACATTTTCAATCTACGCGCGGATACGCGGAAATCATCTGCCACCATTCACGG 223919
781	TGAAGCTTATCGCTTCTGCGGTGATCTGGAAGATTAAATCGGCTTTACCAACCGTCAAAAA 840
223918	TGAAGCTTATCGCTTCTGCGGTGATCTGGAAGATTAAATCGGCTTTACCAACCGTCAAAAA 223860
841	AACGGCGCTTTTATGAGCGCGTTTATTTTCAACTTATTTCCAGATACATCACTCATC 900
223859	AACGGCGCTTTTATGAGCGCGTTTATTTTCAACTTATTTCCAGATACATCACTCATC 223800
901	GTCGGTTGTAATCTTTTACTGCTTT 927
223799	GTCGGTTGTAATCTTTTACTGCTTT 223773
RESULT 10	
AL627284/c	88037 bp DNA linear BCT 16-APR-2005
LOCUS	AL627284
DEFINITION	Salmonella enterica serovar Typhi (Salmonella typhi) strain Ctr18, complete chromosome, segment 20/20.
ACCESSION	AL627284 AL513382
VERSION	AL627284.1 GI:16505618
KEYWORDS	
SOURCE	Salmonella enterica subsp. enterica serovar Typhi
ORGANISM	Salmonella enterica subsp. enterica serovar Typhi
REFERENCE	1 (bases 1 to 88037) Parkhill, J., Dougan, G., James, K.D., Thomson, N.R., Pickard, D., Wain, J., Churcher, C., Mungall, K.L., Bentley, S.D., Holden, M.T., Sebalhia, M., Baker, S., Basham, D., Brooks, K., Chillingworth, T., Connor, P., Cronin, A., Davis, P., Davies, R.M., Dowd, L., White, N., Farrar, J., Felwell, T., Hamlin, N., Haque, A., Hien, T.T., Holroyd, S., Jagels, K., Krogh, A., Larsen, T.S., Leather, S., Mout, S., O'Grady, P., Parry, C., Quail, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S. and Barrell, B.G. Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi Ctr18
AUTHORS	Nature 413 (6858), 848-852 (2001)
PUBMED	11677608
REFERENCE	2 (bases 1 to 88037)
AUTHORS	Parkhill, J.
TITLE	Direct Submission

gene	N-terminal signal sequence"		CDS	signature"	
	/codon_start=1			complement(4840..5853)	
	/transl_table=11			/gene="trpS2"	
	/product="putative exported protein"			/note="synonym: STY4863"	
	/protein_id="CAD03349.1"			complement(4840..5853)	
	/db_xref="GI:16505621"			/gene="trpS2"	
	/db_xref="InterPro:IPR005180"			/EC_number="6.1.1.2"	
	/db_xref="UniProt/TREMBL:Q820Y3"			/note="Similar to Clostridium longisporum	
	/translation="MSTILKPLSLAQNVLMLILIRPGFATEDGAIITMKVTSAYDYLIQT			tryptophanyl-tRNA synthetase TrpS or TrsA SW:SYW_CLOLO	
	RARLMAHADGLVLFGEFEDHARAQNVNLKMPPTVLVFGNPKGDTPLMLAHPELAL			(Q46127) (341 aa) fasta scores: E(): 0, 54.8% id in 330	
gene	DLPFRVLISQOAGDRTLVSYHPAETLQRYGLDADIDQALKLEQLVEKSLH"		CDS	aa, and to Thermotoga maritima tryptophanyl-tRNA	
	complement(3051..3716)			synthetase TrpS SW:SYW_THEME (Q9MYW2) (328 aa) fasta	
	/gene="STY4861"			scores: E(): 0, 36.5% id in 329 aa. Note that the E.coli	
	complement(3051..3716)			trpS ortholog is STY4315"	
	/gene="STY4861"			/codon_start=1	
	/note="Similar to Bacillus subtilis hypothetical protein			/transl_table=11	
	YczH YczH SW:YczH BACSU (O31482) (185 aa) fasta scores:			/product="probable tryptophanyl-tRNA synthetase"	
	E(): 7.4e-16, 31.4% id in 159 aa"			/protein_id="CAD03352.1"	
	/codon_start=1			/db_xref="GI:16505624"	
	/transl_table=11			/db_xref="GOA:Q820Y0"	
misc_feature	/product="hypothetical protein"		CDS	/db_xref="InterPro:IPR001412"	
	/protein_id="CAD03350.1"			/db_xref="InterPro:IPR002305"	
	/db_xref="GI:16505622"			/db_xref="InterPro:IPR002306"	
	/db_xref="InterPro:IPR008262"			/db_xref="UniProt/TREMBL:Q820Y0"	
	/db_xref="UniProt/TREMBL:Q820Y2"			/translation="MNTLFIILTGDRPTGSLHGHVYGSIRQVRVALQDQQVYLIAD	
	/translation="MIKDTVPFAPSVSGSNASSLSRRHRAAIIILHEIYGINAHIQR			LGLDTGNSNPQKIRDNIPQVLADYLAVDIPALTTICLSALPALAELTVLYNNYIV	
	AGHEWTRGFDIYIPALPHPTPPRYEQOEAYRHFSANVGFDPVAVTTLLHDLRAQY			VARVERNPVTKEIQAQGFTRSLPVGFMAYPISQAADITAFKAEVPGVGDQQLPMIBQ	
	ETLIVGVSGATLAWLSAAGLCIDGVCYGSRIQVYLHLAPLCPALVILARYEPAP			TNEIVHKMNSLFSPLVRPCALLSDTGRLPIDGDSAKMSKSLVNTLLISASBETIHR	
	DPMRQILERKPRVQCKMYDARHGFCDAISATFDALSHQVMDVVSFIRIDICANT				
	SPD"				
misc_feature	complement(3375..3404)		CDS	Query Match 77.2%; Score 715.6; DB 1; Length 88037;	
	/gene="STY4861"			Best Local Similarity 92.2%; Pred. No. 2.7e-199;	
	/note="PS00120 Lipases"			Matches 754; Conservative 0; Mismatches 64; Indels 0; Gaps 0;	
	4066..4839			QY 1 GTCATGTTACGCCGATCATGTTAAATTTGCACATCATCAGCAGGTGAGGACGACTTTGT 60	
	/gene="uxuR"			DB 86791 GTCATGTTACGCCGATCATGTTAAATTTGCACATCATCAGCAGGTGAGGACGACTTTGT 86732	
	/note="Synonym: STY4862"			QY 61 ACTTCTCTGTTTCGATTTAGTTGGCAATTTAGTAGCAACATGCAGACCCCGCACATTTCT 120	
	4066..4839			DB 86731 ACTTCTCTGTTTCGATTTAGTTGGCAATTTAGTAGCAACATGCAGACCCCGCACATTTCT 86672	
	/gene="uxuR"			QY 121 TATCGTTGAAGACGAGTTGGTAAACGCAACACGTTGAAAAGTAGTATTTTCGAAGCGGAAGG 180	
	/note="Similar to Escherichia coli uxu operon			DB 86671 TATCGTTGAAGACGAGTTGGTAAACGCAACACGTTGAAAAGTAGTATTTTCGAAGCGGAAGG 86612	
	transcriptional regulator uxuR SW:UXUR ECOLI (P39161) (257			QY 181 CTATGATGTTTTTGAAGCGACAGATGGCGCGAAATGTCATCAGATCCTCTCTGAATATGA 240	
gene	aa) fasta scores: E(): 0, 89.9% id in 257 aa, and to		CDS	DB 86611 CTATGATGTTTTTGAAGCGACAGATGGCGCGAAATGTCATCAGATCCTCTCTGAATATGA 86552	
	Erwinia chrysanthemi exu regulon transcriptional regulator			QY 241 CATCAACCTGGTGATCATGGATATCAATCTGCCGGTAAAGAACGGTCTTCTGTTAGCGCG 300	
	exuR SW:EXUR_ERWCH (Q9X9E0) (259 aa) fasta scores: E(): 0,			DB 86551 CATCAACCTGGTGATCATGGATATCAATCTGCCGGTAAAGAACGGTCTTCTGTTAGCGCG 86492	
	46.9% id in 256 aa. Contains helix-turn-helix motif,			QY 301 TGAAGTGGCGAGCAGCGGAATGTTGCGTTGATGTTCTGATCTGAGCTGCCCGCTGACACGAAGT 360	
	residues 34 to 55, score 975(+2.51 SD).			DB 86491 TGAAGTGGCGAGCAGCGGAATGTTGCGTTGATGTTCTGATCTGAGCTGCCCGCTGACACGAAGT 86432	
	Fasta hit to PDHR_ECOLI (254 aa), 32% identity in 221 aa			QY 361 CGATTAATTTCTCGGCTCGAAATCGGTGAGATGACTACATCACCAACCGTTCAACCC 420	
	overlap			DB 86431 CGATAAATTTCTCGGCTCGAAATCGGTGAGATGACTACATCACCAACCGTTCAACCC 86372	
	Fasta hit to EXUR_ECOLI (258 aa), 48% identity in 255 aa			QY 421 GCGTGAACCTGACGATTTCTGTGACGCAACCTACTGTCCCGTACCATGAATCTGGGTACTGT 480	
	overlap			DB 86371 GCGTGAACCTGACGATTTCTGTGACGCAACCTACTGTCCCGTACCATGAATCTGGGTACTGT 86312	
	Orthologue of E. coli uxuR (UXUR ECOLI); Fasta hit to			QY 481 CAGCGAAGAACGTCGTAGCGTTGAAAGCTACAAGTTTCAATGGTTGGGAACTGGAACATCAA 540	
misc_feature	UXUR_ECOLI (257 aa), 90% identity in 257 aa overlap"		CDS	DB 86311 CAGCGAAGACGTCGTAGCGTTGAAAGCTACAAGTTTCAACCGTTGGGAACTGGAATCAA 86252	
	/codon_start=1			QY 541 CAGCCGTTGTTGATCGGCGCTGATGGCGAGCAGTACAGTCCGCGCGACGAGTTCG 600	
	/transl_table=11			DB 86251 CAGCCGTTGTTGATCGGCGCTGATGGCGAGCAGTACAGTCCGCGCGACGAGTTCG 86192	
	/product="uxu operon transcriptional regulator"				
	/protein_id="CAD03351.1"				
	/db_xref="GI:16505623"				
	/db_xref="GOA:Q820Y1"				
	/db_xref="InterPro:IPR000524"				
	/db_xref="UniProt/TREMBL:Q820Y1"				
	/db_xref="InterPro:IPR011711"				
misc_feature	/translation="MKSNTSQRPYQEVGAMIRDLIVQTPYRPGRLPPERIEARLIN		CDS		
	VTRTVVRALIMLEIKGLIVRRGAGIVYLDSDNNEMEGADVNHCDNAGPELLQAR				
	QLLENIAPFALQATREDIHKWQALOLEERELASSAPGSGESDMOPHLAIARATH				
	NSMLVELFRSQWRENNPMWLQHLSHGLDITLYRKWLGHDHKQILALAIKKDARAKL				
	ANWQHLENVKQLRLLFSNVDDIYFDGYLFESWPLDNVDA"				
	4105..4284				
	/gene="uxuR"				
	/note="Pfam match to entry PF00392 gntR, Bacterial				
	regulatory proteins, gntR family, score 107.20, E-value				
	1.4e-31"				
misc_feature	4162..4236		CDS		
	/gene="uxuR"				
	/note="PS00043 Bacterial regulatory proteins, gntR family				


```

/db_xref="GI:29140249"
/translation="MISAPEPLHAGHILTPFCGVDSDMNLKORAMKNQVGTASRTP
VCCGSDSVLAYISLASAVTNTAPGRFRNMPPDPVIVVLGRVLAVDRSLHGCGVGR
ALVRDAGLRVQAETIGIRGMLVHALSDEAREFYLRVGFEPSPMDPMLWLVLGLDV
HA"
complement(6260..6712)
/locus_tag="t4360"
complement(6260..6712)
/locus_tag="t4360"
/note="corresponds to STY4669 from Accession AL513382:
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="AAO71813.1"
/db_xref="GI:29140250"
/transl_table="NEOCKNDAILHEIKNVSKHIDFBSQANSOGIWLFI STLGCWSV
NIPLLQVTAAILLCIFIPNSKQDMTEKRAFKHIEDIAKIDSLNIGDSRKARLYDL
GLVKRYKAIPVLKTSPIFVICYIFYSISFLVFPSPNLPFRMKLIFFN"
complement(7349..9634)
/locus_tag="t4361"
complement(7349..9634)
/locus_tag="t4361"
/note="corresponds to STY4670 from Accession AL513382:
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAO71814.1"
/db_xref="GI:29140251"
/translation="MVDRYFELAQPFDPVRIWIGNLNFHHQCADQSDQSVQVLR
NDTLRQGIAYVFGPLTDNRNEIFNTRVEKFDHLSHGLHWRNDYKFIINLPEAD
NVDLWAFVQVRRYKNEEOGPDLLRAQMRHRLSKPAFMRWARYNNAMKLSQKS
QLRRPFRKMKRRDRKORETHARNIKFVNENRDIIVERGRHWRCLLRPAELVLMHPEP
TELEFGDEKLVRTALRNLCLDETPEVPTLPALALQCSKYRHSHTVLYAACLEILRA
EGNLCVNIELTALRNIHRGHSVSKERDALQTEIDRLIFPYSSEAETKLQYVE
PQLTCPPHEPIGLVRSQEVFRHRSRAKLSIEWLRRFPDVSLSNVDTLFEIAAQYGDRE
NLKEIIARCSMMSGNWPNKEDIERKIFWLVRBFYFMENADTYWMLKSDKN
LHIFYEGRMNSHRWAPPELASIKVEAILDAFT EHWPRVDLPDSMGSDPKEKAY
RFLTDLIWSINDTDDAI PVLDRLLNDPRFTNLHKEIQSI VADQIRKKALRDEPPT
PDEIVORLDCSVTVTEGLRQLVLQELHDFOKAIDGGEFNSADRPYKEXERLDEKST
EIIARLSRLQPOGIATPHEQLKQNRSDFTASKLIGRRRLLVITVSKQWHLRG
SAASQLVDYSIHDPDQOQGLFVIFWGESETVAGRKTGHGKTAQELKLSIEAVLPA
complement(9715..11388)
/locus_tag="t4362"
complement(9715..11388)
/note="corresponds to STY4671 from Accession AL513382:
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="AAO71815.1"
/db_xref="GI:29140252"
/translation="MNSIFYLERHFTHGDRYVTELELAASHVVVLAEPGGKTELM
KSLARKLNTVLNAGVAVAGADKENSALVIDAVEVARIDQSGHKLALALARTSKPT
RVIMSRSEWGQACTSIFPEKFLAGFSPMVLKKEFDQDOEQHAI PKHHAPEFPAFOT
EYTRSLDMLLPNPFKMPFTDYLSEGRFPADKRSIFALAVERLAKENVNDIPKAPV
SLSVAQKISFAEVTAKLLSGAGVSTIDVTANRWYPMPLPALFSSAACVDYILSTQL
FKPGDKDQHPVRVHKIYVAEYCAADYLIKRIADPDVLDVLPKCLPVIAPNGIARDELRG
LIGWMAALGNRSVQSGSIIELDAYAVLANGDPSQLRSKRQLLYOLKIEADPYFRR
SDFMRFAGAGFTQDVVVEEIKPLLTMGNDGHLRLIILELLADSPVNSQLVPELSQLT
LNPNSRHRTLASCLNLIKEDYICIGALAVLIPEASNISLNIAANVLEVTGPEKFDH
RYLSGFPFRVCN396
11935..12396
/locus_tag="t4363"
11935..12396
/locus_tag="t4363"
/note="corresponds to STY4672 from Accession AL513382:
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
```

```

Query Match 77.2%; Score 715.6; DB 1; Length 284233;
Best Local Similarity 92.2%; Pred. No. 3.3e-199;
Matches 754; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 1 GTCATGTTACGCGCATCATGTTAAATTTGACAGATGTCATCAGGACAGGTTCAGGACCTTTGT 60
DB 282987 GTCATGTTACGCGCATCATGTTAAATTTGCGACATGTCATCAGGACAGGTTCAGGACCTTTGT 282928

QY 61 ACTTCCCTGTTTCGATTTAGTTGGCAATTTAGGTAGCAAAACATGCAAGACCCCGCACATCT 120
DB 282927 ACTTCCCTGTTTCGATTTAGTTGGCAATTTAGGTAGCAAAACATGCAAGACCCCGCACATCT 282868

QY 121 TATCGTTGAAGACGAGTTGGTAAACGCAACACCGTGTGAAAAAGTATTTTCGAAGCGCGAAGG 180
DB 282867 TATCGTTGAAGACGAGTTGGTAAACGCAACACCGTGTGAAAAAGTATTTTCGAAGCGCGAAGG 282808

QY 191 CTATGATGTTTTCGAAGCGACAGATGGCGCGGAAATGTCATCAGATCTCTCTGTAATATGA 240
DB 282807 CTATGATGTTTTCGAAGCGACAGATGGCGCGGAAATGTCATCAGATCTCTCTGTAATATGA 282748

QY 241 CATCAACCTGCTGATCATGGATATCAATCTCCCGGTAAAGAACGCTCTCTGTTAGCGCG 300
DB 282747 CATCAACCTGCTGATCATGGATATCAATCTCCCGGTAAAGAACGCTCTCTGTTAGCGCG 282688

QY 301 TGAACCTGCGGACGAGCGGAATGTTGCGTTGATGTTCTGACTGCGCGGTGACCAACGAAAGT 360
DB 282687 TGAACCTGCGGACGAGCGGTATGTTGCGTGTGTTCTGAGCGGCGCGGATACGAAAGT 282628

QY 361 CGATAAATTTCTCGGCTCGAAATCGGTGCGAGATGACTATACATCAACAAACCGTTCAACCC 420
DB 282627 CGATAAATTTCTCGGCTCGAAATCGGTGCGGAGATGATTAACATCAACAAACCGTTTAAATCC 282568

QY 421 GCGTGAACCTGACGATTCGTGACGCAACCTACTGTCGCGTACCATGAATCTGGGTACTGT 480
DB 282567 ACGGAAATGACCATTCGTCGCGCGGTAACTGCTGTCGCGTACCATGAACCTGGGTACTGT 282508

QY 481 CAGCGAAGAACCTGCTAGCTGTTGAAAGCTACAAGTTCATGTTGGGAACTGGACATCAA 540
DB 282507 CAGCGAAGAACCTGCTAGCTGTTGAAAGCTATAAGTTCACGTTGGGAACTGGATATCAA 282448

QY 541 CAGCGGCTGTTGATCGGCGCTGATGCGGACAGATGACAGCTGCGCGCGGAGGAGTTCG 600
DB 282447 CAGCGGTTCACTGATCGGCGCGGACGAGCAGTATAAACTACCGCGCAGCGAGTTCG 282388

QY 601 CGCCATGCTTCACTTCTGTGAAAACCCAGGCAAAATTCAGTCCCGTCTGAACTGCTGAA 660
DB 282387 CGCCATGCTTCACTTCTGTGAAAACCCCTGGCAAAATTCAGTCTCGCGCGAGCTGCTGAA 282328

QY 661 GAAAATGACCGCGCGTGTAGCTGAAAACCCGACGACCGTACTGTAGACGTGACGATCCGCG 720
DB 282327 AAAAAATGACCGGACGCGGAGCTGAAAGCCGATGACCGTACTGTGAGCTCACCATTCGTCG 282268

QY 721 TATTCGTAAACATTTGCAATCTACCGCGGATACGCGCGGAAATCATCGCCACCATTCACGG 780
DB 282267 TATTCGTAAAGCATTTGCAATCCACCGCGGATACGCGCGGAAATCATCGCCACCATTCACGG 282208

QY 781 TGAAGGTTATCGTCTTCGCGGTGATCTGGAAGATTAAT 818
DB 282207 CGAAGGTTATCGTCTTCGCGCGGACCTGCGAGATTAAT 282170

RESULT 12
LOCUS AE008916/c
DEFINITION Salmonella typhimurium LT2, section 220 of the complete genome.
ACCESSION AE008916
VERSION AE008916.1 GI:16423153
KEYWORDS
SOURCE Salmonella typhimurium LT2
ORGANISM Salmonella typhimurium LT2
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
```

Enterobacteriaceae; Salmonella.

REFERENCE
AUTHORS
1 (bases 1 to 13852)
McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W.,
Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F.,
Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A.,
Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W.,
Stoneking, T., Nhan, M., Waterston, R. and Wilson, R.K.
Complete genome sequence of *Salmonella enterica* serovar Typhimurium
LT2
Nature 413 (6858), 852-856 (2001)
11677609
2 (bases 1 to 13852)
The Salmonella typhimurium Genome Sequencing Project
Direct Submission
Submitted (29-MAR-2001) Genome Sequencing Center, Department of
Genetics, Washington University School of Medicine, 4444 Forest
Park Boulevard, St. Louis, MO 63108, USA
Supported by NIH grant 5U 01 AI43283

Coding sequences below are predicted from manually evaluated
computer analysis, using similarity information and the programs;
GLIMMER; <http://www.tigr.org/softlab/glimmer/glimmer.html> and
GeneMark; <http://opal.biology.gatech.edu/GeneMark/>

EC numbers were kindly provided by Junko Yabuzaki and the Kyoto
Encyclopedia of Genes and Genomes; <http://www.genome.ad.jp/kegg/>,
and Pedro Romero and Peter Karp at EcoCyc;
<http://ecocyc.PangeaSystems.com/ecocyc/>

The analyses of ribosome binding sites and promoter binding sites
were kindly provided by Heladia Salgado, Julio Collado-Vides and
RegulonDB;
http://kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistries or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one m13 subclone.

FEATURES

Source

1. 13852
/organism="Salmonella typhimurium LT2"
/mol_type="genomic DNA"
/strain="LT2; SGSC 1412; ATCC 700720"
/db_xref="ATCC:700720"
/db_xref="taxon:99287"
/note="LT2"
90..640
/gene="creA"
/note="synonym: STM4587"
90..98
/gene="creA"
/note="putative -35 signal for creA;
RegulonDB:STMLTH004682"

gene

-35_signal

-10_signal

117..125
/gene="creA"
RegulonDB:STMLTH004682"

RBS

156..161
/gene="creA"
/note="putative RBS for creA; RegulonDB:STMSIH004447"
167..640

CDS

/gene="creA"
/note="similar to E. coli orf, hypothetical protein
(AAC77350.1); Blastp hit to AAC77350.1 (157 aa), 89%
identity in aa 1 - 157"
/codon_start=1
/transl_table=1
/product="putative periplasmic protein"
/protein_id="AAL23402.1"
/db_xref="GI:16423154"

/translation="MKYKSLVLFSLILLGQSARAEQIGSDVTFFKMGPGDKHIVVEA
FDDPDKNVTCYSRAKTGGIKGGLAEDTSDAISCQOQVPIELSDKIKNGKAQGE
VFKKRLSLFKSLQVVRFDKRNKTLAYLAYSQVVDGSPKNAISAVPVMWRQ"
641..1342
/gene="creB"
/note="synonym: STM4588"
641..646
/gene="creB"
/note="putative RBS for creB; RegulonDB:STMSIH004448"
653..1342
/gene="creB"
/note="OmpR family; similar to E. coli catabolic
regulation response regulator (AAC77351.1); Blastp hit to
AAC77351.1 (229 aa), 83% identity in aa 1 - 229"
/codon_start=1
/transl_table=1
/product="response regulator in two-component regulatory
system with CreC"
/protein_id="AAL23403.1"
/db_xref="GI:16423155"
/translation="MQSPQWLVDEQGIADTLIYTLQEGTVELFARGLPALREKAR
QQRDAVLDVGLPDISGFCQLLERHPALPIFLFARSDEVDRLLGLEAGADYY
AKPFSREVSARVETLLRRVKKFAAPSPVVFHFLNEPAQAIAWFTPLSLTYEF
LLKTLKLSPERVYSROQLMDIVMSDAQETDRTVDTHIKTLRAKLAINPELSPINT
HRGMGYSLRSV"
1330..2766
/gene="creC"
/note="synonym: STM4589"
1330..1335
/gene="creB"
/note="putative RBS for creC; RegulonDB:STMSIH004449"
1342..2766
/gene="creC"
/EC_number="2.7.3.-"
/note="senses catabolite repression; similar to E. coli
catabolite repression sensor kinase for PhoB; alternative
sensor for pho regulon (AAC77352.1); Blastp hit to
AAC77352.1 (474 aa), 80% identity in aa 1 - 474"
/codon_start=1
/transl_table=1
/product="sensory kinase (alternative) in two-component
regulatory system with CreB (or alternatively PhoB)"
/protein_id="AAL23404.1"
/db_xref="GI:16423156"
/translation="MRIGMRLLGLFLIVAAVPLSIFVQEIKPGVRRATEGLID
TATLLAARDLDDGMPDGLAKAFQLOHREFRANISGIVKVRNETHYIMDAQG
KVLFDSENKAVGDYSRWNDVQLTRGQYGARSTAKDAADPNVTMYAIPISDGR
IGLVTVGKNAAPVTKRSRRILWASAVLLGALAIAGVMWIMNRISARLTGRAD
SVTENRPLALPGSLERLKLQALLESNRILEGKNIEQVYVALTHELKSPLAARIG
AAETLRGPPADVTRPTENTLAQNRMAQVETLLROALNRQDIPLAPVAVDEL
TQSEARSILQAAKTLTLTPSSIVVADAELLAQALGNVLDNAIDFTPENGVLTLS
AQPMEGRKAILQVDSGCCIPDFALPRIDRFYSLPRENGRKSSGLGLAFVSEARLLN
GEVTLGNRPEGGLASLTLRHFT"
2812..4173
/gene="creB"
/note="synonym: STM4590"
2812..2817
/gene="creB"
/note="putative RBS for creD; RegulonDB:STMSIH004450"
2824..4173
/gene="creB"
/note="similar to E. coli tolerance to colicin E2
(AAC77353.1); Blastp hit to AAC77353.1 (450 aa), 73%
identity in aa 1 - 450"
/codon_start=1
/transl_table=1
/product="tolerance to colicin E2 protein"
/protein_id="AAL23405.1"
/db_xref="GI:16423157"
/translation="WLKSPLLWKMTLGGMILLILIPLMYVRHTIVERADYSHVEAA
IRQSTSGPKVGVPLVAIPVTELYTVLEEEVKYKRSILYFLWPESLVGNQNVFA
RKIGTYQGVHWTMDIAKEEDVARLHNLNPNITLGPFFIVGVGDARGISAVKAPQ
VNGETLTVEPGTGLPESREGIHIPLPDSQWATNRNLTLDMSLNLGTSFSLPVGRRS

gene

RBS

CDS

AE017220_34	3400001	3510000
AE017220_35	3500001	3610000
AE017220_36	3600001	3710000
AE017220_37	3700001	3810000
AE017220_38	3800001	3910000
AE017220_39	3900001	4010000
AE017220_40	4000001	4110000
AE017220_41	4100001	4210000
AE017220_42	4200001	4310000
AE017220_43	4300001	4410000
AE017220_44	4400001	4510000
AE017220_45	4500001	4610000
AE017220_46	4600001	4710000
AE017220_47	4700001	4755700
Continuation (48 of 48) of AE017220 from base 4700001 (AE017220 Salmonella enterica subsp. enterica)		
Query Match 77.0%; Score 714; DB 1; Length 55700;		
Best Local Similarity 92.1%; Pred. No. 7.3e-199;		
Matches 753; Conservative 0; Mismatches 65; Indels 0; Gaps 0;		
QY	1	GTCATGTTACGGCGATCATGTTAAATTCGACGATCATCAGCAGTTCAGGACTTTTGT 60
DB	54454	GTCAATGTTACGGCGATCATGTTAAATTCGACATCATCAGCAGTTCAGGACTTTTGT 54395
QY	61	ACTTCTGTTTCGATTTAGTTGGCAATTTAGGTAGCAACATGCAAGCCCGCACATTCT 120
DB	54394	ACTTCTGTTTCGATTTAGTTGGCAATTTAGGTAGCAACATGCAAGCCCGCACATTCT 54335
QY	121	TATCGTTGAAGCAGTGGTTGAACACGCAACAGTGTGAAGATATTTTGAAGCGGAAG 180
DB	54334	TATCGTTGAAGCAGTGGTTGAACACGCAACAGTGTGAAGATATTTTGAAGCGGAAG 54275
QY	181	CTATGATGTTTTCGAAGCAGATGCGCGGAATGCATCAGATCCTCTGATATATGA 240
DB	54274	CTATGATGTTTTCGAAGCAGATGCGCGGAATGCATCAGATCCTCTGATATATGA 54215
QY	241	CATCAACCTGGTGATCATGATATCAATCTGCGGTGAAGACGCTTCTGTTAGCGG 300
DB	54214	CATCAACCTGGTGATCATGATATCAATCTGCGGTGAAGACGCTTCTGTTAGCGG 54155
QY	301	TGAACCTGCGCAGCAGGGAATGTTGCGTTGATGTTCTGATGCGCGGTGACAAACGAGT 360
DB	54154	TGAACCTGCGTGAACAGGCTAACGTTGCGCTGATGTTCTGACGCGCGCGATACGAAAT 54095
QY	361	CGATTAATTTCTCGGCTCGAATCGGTGCAGATGACTACATCAACCAACCGTTCAACCC 420
DB	54094	CGACAAATTTCTCGGCTCGAATCGGCGCGATGATTAATCATCAACCAACCAATTAACCC 54035
QY	421	GGGTGAACCTGACGATTCGTGACGCAACCTACTGTCGGTACCATGATCTGGGTACTGT 480
DB	54034	ACGCGNATTTGACATCCGTGCGGCAACCTGTGTCGGTACCATGATCTGGGTACTGT 53975
QY	481	CAGCGAAGACCTGCTAGCGTTGAAAGCTACAAGTTCAATGTTGGGAACCTGACATCAA 540
DB	53974	CAGCGAAGACCTGCTAGCGTTGAAAGCTACAAGTTCAACGTTGGGAACCTGATCAA 53915
QY	541	CAGCGTTGCTGATCGGCGCTGATGCGGACAGTACAAGTTCGCGCGCAGCGAGTTCCG 600
DB	53914	CAGTGTTCACGTGATCGGCGCGGAGGAGGATTAACATACCGCGCAGCGAGTTCCG 53855
QY	601	CGCCATGCTTCACTTCTGTGAAACCCAGGCAAAATTCAGTCCCGTGTGAATCTGTA 660
DB	53854	CGCGATGCTTCACTTCTGTGAAACCCCTGCGAATATTCAGTCCCGCGCAGCTGTGNA 53795
QY	661	GAAATGACCGCGCTGAGCTGAAACCGCACGACCTGCTAGACGTGAGACGATCCGCG 720
DB	53794	GAAATGACCGCGCTGAGCTGAAAGCGCATGACCGTACTGTTGACGTTACCATTCGTCG 53735
QY	721	TATTCGTAAATTTTCAATCTACCGCGATACCGCGGAATCATCGCCACCATTCAGG 780
DB	53734	TATTCGTAAAGATTTTCAATCTACCGCGATACCGCGGAATCATCGCCACCATTCAGG 53675
QY	781	TGAAGGTTATCGCTTCTCGGCTGATCTGGAAGATTAAT 818

Db	53674	CGAAGTTATCGCTTCTGCGCGACCTGCAGGATTAAT 53637
RESULT 14		
CP000026_45/c		
WPCOMMENT		
Sequence split into 46 fragments LOCUS CP000026 Accession CP000026		
Fragment Name Begin End		
CP000026_00	1	110000
CP000026_01	100001	210000
CP000026_02	200001	310000
CP000026_03	300001	410000
CP000026_04	400001	510000
CP000026_05	500001	610000
CP000026_06	600001	710000
CP000026_07	700001	810000
CP000026_08	800001	910000
CP000026_09	900001	1010000
CP000026_10	1000001	1110000
CP000026_11	1100001	1210000
CP000026_12	1200001	1310000
CP000026_13	1300001	1410000
CP000026_14	1400001	1510000
CP000026_15	1500001	1610000
CP000026_16	1600001	1710000
CP000026_17	1700001	1810000
CP000026_18	1800001	1910000
CP000026_19	1900001	2010000
CP000026_20	2000001	2110000
CP000026_21	2100001	2210000
CP000026_22	2200001	2310000
CP000026_23	2300001	2410000
CP000026_24	2400001	2510000
CP000026_25	2500001	2610000
CP000026_26	2600001	2710000
CP000026_27	2700001	2810000
CP000026_28	2800001	2910000
CP000026_29	2900001	3010000
CP000026_30	3000001	3110000
CP000026_31	3100001	3210000
CP000026_32	3200001	3310000
CP000026_33	3300001	3410000
CP000026_34	3400001	3510000
CP000026_35	3500001	3610000
CP000026_36	3600001	3710000
CP000026_37	3700001	3810000
CP000026_38	3800001	3910000
CP000026_39	3900001	4010000
CP000026_40	4000001	4110000
CP000026_41	4100001	4210000
CP000026_42	4200001	4310000
CP000026_43	4300001	4410000
CP000026_44	4400001	4510000
CP000026_45	4500001	4585229
Continuation (46 of 46) of CP000026 from base 4500001 (CP000026 Salmonella enterica subsp. enterica)		
Query Match 76.9%; Score 712.4; DB 1; Length 85229;		
Best Local Similarity 91.9%; Pred. NO. 2.3e-198;		
Matches 752; Conservative 0; Mismatches 66; Indels 0; Gaps 0;		
QY	1	GTCATGTTACCGCGATCATGTTAAATTCGACGATCATCAGCAGTTCAGGACTTTTGT 60
DB	83983	GTCAATGTTACCGCGATCATGTTAAATTCGACATCATCAGCAGTTCAGGACTTTTGT 83924
QY	61	ACTTCTGTTTCGATTTAGTTGGCAATTTAGGTAGCAACATGCAAGCCCGCACATTCT 120
DB	83923	ACTTCTGTTTCGATTTAGTTGGCAATTTAGGTAGCAACATGCAAGCCCGCACATTCT 83864
QY	121	TATCGTTGAAGCAGTGGTTGAACACGCAACAGTGTGAAGATATTTTGAAGCGGAAG 180
DB	83863	TATCGTTGAAGCAGTGGTTGAACACGCAACAGTGTGAAGATATTTTGAAGCGGAAG 83804

Qy	181	CTATGATGTTTTG	AAGCGACAGATCGCGGGAAATGCAATCAGATCCTCTCTGAATATGA	240
Db	83803	CTATGATGTTATTC	GAAAGCGACAGATCGCGGGAAATGCAATCAGATCCTCTCTGAATATGA	83744
Qy	241	CATCAACCTGGTGAT	CATGAGATCAATCTGCCGGTAAAGCGGTCTCTGTTAGCGCG	300
Db	83743	CATCAACCTGGTGAT	CATGAGATCAATCTGCCAGGAAAAACGGTCTCTGTTAGCGCG	83684
Qy	301	TGAATCGCGGAGCG	CGAATGTTCTGGTTGATGTTCTTGACTGCGCGGTGACACAAAGT	360
Db	83683	TGAATCGGTGAAAC	GAGCTAAACGTTGCGGTGATGTTCTTCTGACGGCGCGGATAACAAAGT	83624
Qy	361	CGATAAATTTCTCG	GCGCTCGAATCGGTGCAGATGACTACATCAACCAACCGTTCAACCC	420
Db	83623	CGACAAATTTCTCG	GCGCTCGAATTCGGCGGAGATTAATCATCAACAAACGGTTTAAACC	83564
Qy	421	GCGTGAATCAGCA	ATTCGTGCACGCAACCTACTGTCTCCGTACCATGAATCTCGGGTACTGT	480
Db	83563	ACGCGAATTGACCA	TTCGTTGCCGCCCAACCTGCTGTCTCCGTACCATGAACCTGGGTACCGT	83504
Qy	481	CAGCGAAGAACGTC	GTAGCGTTGAAAAGCTACAAGTTCAATGTTGGGAACCTGGAACATCAA	540
Db	83503	CAGCGAAGACGTC	GTAGCGTTGAAAAGCTACAAGTTCAACGGTTGGGAACCTGGAATATCAA	83444
Qy	541	CAGCGGTTCTGTAT	CGGGCCCTGATCGCAGCAGCAGTACAGCTCCGCGCAGCGAGCTCCG	600
Db	83443	CAGTGTGTTACTGAT	CGGCCCGGACGGCAGCAGTATAAATACCGCGCAGCGAGTTCGG	83384
Qy	601	CGCCATGCTTTCACT	TCTGTGAAAACCCAGGCAAAATTCAGTCCCGTGTGTAACCTCTGAA	660
Db	83383	CGCGATGCTGCACTT	CTGCGAAAACCTTCGCAAAATTCAGTTCGCGCGCAGAGCTGTTGAA	83324
Qy	661	GAAATTCACGCGCG	GTAGCTGAAAACCGCACGACCGTACTGTGAGGTGACGATCCGCGCG	720
Db	83323	GAAATTCACGCGCG	GAGCTGAAGCCGATACGCGGTACTGTGTGACGTTACCATTCGTGCG	83264
Qy	721	TATTTCTGTAAACAT	TTTCGAATCTACGCGCGGATACGCGGAAATTCATCGCCACCATTCACGG	780
Db	83263	TATTTCTGTAAACAT	TTTCGAATTCACGCGCGGATACGCGGAAATTCATCGCCACCATTCACGG	83204
Qy	781	TGAAGGTTATCGCTT	CTCGGGTGATCTGGAAGATTAAAT	818
Db	83203	CGAAGGTTATCGCTT	CTCGGGCACCTGACGAGTTAAAT	83166

RESULT 15	AY832936	AY832936	622 bp	DNA	linear	BCT 26-APR-2005
LOCUS	AY832936	Escherichia coli isolate 044 aerobic respirator control protein A (arca) gene, partial cds.				
DEFINITION	AY832936	AY832936.1	GI:56463110			
ACCESSION	AY832936	Escherichia coli				
VERSION	AY832936	Escherichia coli				
KEYWORDS	AY832936.1	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.				
SOURCE	AY832936.1	1 (bases 1 to 622)				
ORGANISM	AY832936.1	Nemoy, L.L., Kotetishvili, M., Tigno, J., Keefer-Norris, A., Harris, A.D., Perencevich, E.N., Johnson, J.A., Torpey, D., Sulakvelidze, A., Morris, J.G. Jr. and Stine, O.C.				
REFERENCE	AY832936.1	Multilocus sequence typing versus pulsed-field gel electrophoresis for characterization of extended-spectrum beta-lactamase-producing Escherichia coli isolates				
AUTHORS	AY832936.1	J. Clin. Microbiol. 43 (4), 1776-1781 (2005)				
TITLE	AY832936.1	15814998				
JOURNAL	AY832936.1	2 (bases 1 to 622)				
PUBLISHED	AY832936.1	Nemoy, L.L., Kotetishvili, M., Tigno, J., Keefer-Norris, A., Harris, A.D., Perencevich, E.N., Johnson, J.A., Torpey, D., Sulakvelidze, A., Morris, J.G. Jr. and Stine, O.C.				
REFERENCES	AY832936.1	Direct Submission				
AUTHORS	AY832936.1	Submitted (18-NOV-2004) Epidemiology and Preventive Medicine, University of Maryland School of Medicine, 10 South Pine Street,				
TITLE	AY832936.1					
JOURNAL	AY832936.1					

FEATURES	source	location/Qualifiers
gene	gene	1..622
CDS	CDS	1..622
		/organism="Escherichia coli"
		/mol_type="genomic DNA"
		/isolate="044"
		/db_xref="taxon:562"
		<1..>622
		/gene="arcA"
		<1..>622
		/genes="arcA"
		/codon_start=1
		/transl_table=11
		/product="aerobic respirator control protein A"
		/protein_id="AAV91697.1"
		/db_xref="GI:56463111"
		/translation="KSIIFAEGYDVFPEATDGAEMHQLSEYDINLVIMDINLPKGNKL MLARELRQANVALMELKDRONEVDKILGLEICADDYITKPFNPRLRTTARRNLRT LNLGTVSERRRSVESYKFNQWELDINSRLIGPDGBOYKLPSEPRAMLHFCENPGKI QSRBELKKMTGRELKPHDRTVDVTRIRKHPESPDPTEIIATIH"
ORIGIN		
	Query Match	66.4%; Score 615.6; DB 1; Length 622;
	Best Local Similarity	99.4%; Pred. No. 3.7e-170;
	Matches	618; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY	158	AAAAGTATTTCGAAGCGGAAGCTATGATGTTTCGAAAGCGACAGATGCGCGGAATG 217
Db	1	AAAAGTATTTCGAAGCGGAAGCTATGATGTTTCGAAAGCGACAGATGCGCGGAATG 60
QY	218	CATCAGATCCTCTCTGAATATGACATCAACCTGGTGATCATGTGATATCAATCTGCCGGT 277
Db	61	CATCAGATCCTCTCTGAATATGACATCAACCTGGTGATCATGTGATATCAATCTGCCGGT 120
QY	278	AAGAACGGTCTTCTGTAGCGGTGAACCTGCGGAGCAGCGGAACTGTTCGGTTGATGTTTC 337
Db	121	AAAAACGGTCTTCTGTAGCGGTGAACCTGCGGAGCAGCGGAACTGTTCGGTTGATGTTTC 180
QY	338	CTGACTGGCGGTGCACAAACGAAGTCGATAAAATTTCTCGGCTCGAAATCGGTGCAGATGAC 397
Db	181	CTGACTGGCGGTGCACAAACGAAGTCGATAAAATTTCTCGGCTCGAAATCGGTGCAGATGAC 240
QY	398	TACATCACAAACCGTTTCAACCGCGTGAACTGACGATTCGTGCAACGAACTTCTGTGTC 457
Db	241	TACATCACAAACCGTTTCAACCGCGTGAACTGACGATTCGTGCAACGAACTTCTGTGTC 300
QY	458	CGTACCATGAATCTGGGTACTGTGACGGAAGACGTCGTAGCGTTCAAGAGCTACAAGTTC 517
Db	301	CGTACCATGAATCTGGGTACTGTGACGGAAGACGTCGTAGCGTTCAAGAGCTACAAGTTC 360
QY	518	AATGGTTGGGAATCGACATCAACAGCCGTTCTGGTATCGGCCCTGATGCGGAGCAGTAC 577
Db	361	AATGGTTGGGAATCGACATCAACAGCCGTTCTGGTATCGGCCCTGATGCGGAGCAGTAC 420
QY	578	AAGCTCGCGGACGAGTTCCGCGCATGCTTCACTTCTGTGAAAACCCAGCAAAATTT 637
Db	421	AAGCTCGCGGACGAGTTCCGCGCATGCTTCACTTCTGTGAAAACCCAGCAAAATTT 480
QY	638	CAGTCCCGTCTGAACTGTGTGAAGAAAATGACCGCGCGTGAAGCTGAAACCGCACGCCGT 697
Db	481	CAGTCCCGTCTGAACTGTGTGAAGAAAATGACCGCGCGTGAAGCTGAAACCGCACGCCGT 540
QY	698	ACTGTAGACGTGACGATCGCCGTATTCGTAAACATTTTGAATCTAGCCCGGATACGCCG 757
Db	541	ACTGTAGACGTGACGATCGCCGTATTCGTAAACATTTTGAATCTAGCCCGGATACGCCG 600
QY	758	GAAATCATCGCCACCATTTCAAG 779
Db	601	GAAATCATCGCCACCATTTCAAG 622

Search completed: February 13, 2006, 15:00:59
Job time : 4714 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2006, 13:28:02 ; Search time 601 Seconds
(without alignments)
10279.817 Million cell updates/sec

Title: US-10-613-990A-31
Perfect score: 927
Sequence: 1 gtcgtgtaacccgatcatg.....gtaactttcttactgggttt 927

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001as:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2002bs:.*
8: Geneseqn2003as:.*
9: Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004as:.*
13: Geneseqn2004bs:.*
14: Geneseqn2005s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES											
Result	No.	Score	Query Match	Length	DB ID	Description					
c	1	927	100.0	927	12	ADJ38402	Escherich				
	2	925.8	99.9	1732	6	ABL90421	Human pol				
	3	577.2	62.3	726	11	ACH96017	Klebsiell				
	4	510.6	55.1	110000	10	ACF67367	Continuati				
	5	510.6	55.1	110000	10	ACF65386	Phototrab				
c	6	508.6	54.9	759	12	ADJ38390	Pantosea a				
	7	493	53.2	717	10	ACF69844	Phototrab				
	8	450.8	48.6	726	5	ADF01748	Bacterial				
	9	447.4	48.3	1048	5	AA592939	DNA encod				
	10	320.6	34.6	191996	13	ADT05647	Haemophil				
c	11	319.4	34.5	110000	2	AT42063	Continuati				
	12	316.2	34.1	711	4	AA553409	Haemophil				
	13	316.2	34.1	711	8	ACA34287	Prokaryot				
	14	232.8	25.1	1110	5	AA582422	DNA encod				
	15	161.6	17.4	2125	13	ADT05255	Haemophil				
c	16	153.6	16.6	2227	13	ADT05194	Haemophil				
	17	137.4	14.8	729	6	ABQ90117	M. capsul				
	18	126	13.6	837	11	ABD17813	Pseudomon				
	19	126	13.6	2430	11	ABD17567	Pseudomon				

c	20	126	13.6	3675	11	ABD17705	Abd17705 Pseudomon
	21	120.6	13.0	990	11	ABD12913	Abd12913 Pseudomon
	22	120.6	13.0	2166	11	ABD13118	Abd13118 Pseudomon
	23	113.8	12.3	2004	11	ABD13261	Abd13261 Pseudomon
	24	107.4	11.6	1018	3	AAA38554	AAA38554 Actinobac
c	25	106.8	11.5	11275	6	AAH43717	Aah43717 E. coli g
	26	103.8	11.2	714	8	ACA47574	ACA47574 Prokaryot
	27	101.4	10.9	705	8	ACA48118	ACA48118 Prokaryot
	28	100	10.8	100	8	ACD69947	ACd69947 E. coli K
	29	100	10.8	100	8	ACD69946	ACd69946 E. coli K
	30	100	10.8	100	8	ACD69948	ACd69948 E. coli K
	31	99	10.7	732	8	ACA28627	ACA28627 Prokaryot
	32	98.6	10.6	110000	6	AAH03041	Continuati
	33	98	10.6	714	8	ACA36700	Continuati (4 of
c	34	97.2	10.5	762	11	ABD07703	Abd07703 Pseudomon
	35	96.8	10.4	349980	6	ABQ81844	Abq81844 Bifidobac
	36	96.2	10.4	705	6	ABN90763	ABn90763 Staphyloc
	37	96.2	10.4	705	13	ADS03841	AdS03841 Staphyloc
	38	96.2	10.4	721	8	ACA46767	ACA46767 Prokaryot
	39	95	10.2	110000	6	ABQ69245	Continuati (4 of
	40	95	10.2	110000	6	ABQ67197	Continuati (3 of
	41	94.6	10.2	702	4	AAH52543	Aah52543 S. epider
	42	94.6	10.2	3389	4	AAH54550	Aah54550 S. epider
	43	93	10.0	699	4	AA552039	AA552039 Staphyloc
	44	93	10.0	702	3	AAA95530	AAa95530 S. aureus
	45	93	10.0	702	13	ADZ74065	Adz74065 Staphyloc

ALIGNMENTS

RESULT 1	
ADJ38402	
ID	ADJ38402 standard; DNA; 927 BP.
XX	
AC	ADJ38402;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	Escherichia coli arca (aerobic respiration control) gene SeqID31.
XX	
KW	gamma-proteobacterium; Arca; aerobic respiration control;
KW	Arca protein function; chromosomal arca gene; fermentation; L-amino acid;
KW	L-lysine; L-glutamic acid; gene; ds.
XX	
OS	Escherichia coli.
XX	
Key	Location/Qualifiers
FT	101..817
CDS	/*tag= a
FT	/product= "Escherichia coli arca protein"
FT	
XX	
PN	EP1382686-A1.
XX	
XX	21-JAN-2004.
XX	
XX	11-JUL-2003; 2003EP-00015911.
XX	
XX	12-JUL-2002; 2002JP-00203764.
XX	
XX	(AJIN) AJINOMOTO CO INC.
XX	
XX	Ishikawa Y, Imaizumi A, Matsui K, Kojima H;
PI	
XX	WPI; 2004-192975/19.
DR	P-PSDB; ADJ38403.
DR	
XX	
PT	New bacterium useful for producing a target substance (especially L-
PT	lysine, L-glutamic acid or another L-amino acid) is modified so that the
PT	Arca protein does not normally function.
XX	
XX	Claim 6; SEQ ID NO 31; 38pp; English.
PS	
XX	

CC This invention relates to a novel gamma-proteobacterium able to produce a
CC target substance such as an amino acid which is modified so that an Arca
CC (aerobic respiration control) protein does not normally function. The
CC disruption to Arca protein function is preferably due to disruption of a
CC chromosomal arca gene. The bacterium and method are useful in producing a
CC target substance by fermentation. The target substance is preferably an L
CC -amino acid, especially L-lysine or L-glutamic acid. Production of the
CC target substance using the modified bacterium is more efficient. The
CC present sequence is that of a PCR primer which was used for amplification
CC of the P ananatis arca gene in the exemplification of the invention.
XX

SQ Sequence 927 BP; 236 A; 228 C; 225 G; 238 T; 0 U; 0 Other;

Query Match	100.0%;	Score 927;	DB 12;	Length 927;
Best Local Similarity	100.0%;	Pred. No. 1.4e-237;		
Matches 927;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GTCTGTTTACGCGCATCATGTTAAATTTGCGATGATCATCAGCAGGTGAGGACTTTTGT	60	
Db	1	GTCTGTTTACGCGCATCATGTTAAATTTGCGATGATCATCAGCAGGTGAGGACTTTTGT	60	
Qy	61	ACTTCTGTTGATTTAGTTGGCAATTTAGGTAGCAAAACATGCGAGCCCGCACATTTCT	120	
Db	61	ACTTCTGTTGATTTAGTTGGCAATTTAGGTAGCAAAACATGCGAGCCCGCACATTTCT	120	
Qy	121	TATCTTTGAGACGAGTTGGTAAACGCAACACGTTGAAAGTATTTTCGAGCGGAGG	180	
Db	121	TATCTTTGAGACGAGTTGGTAAACGCAACACGTTGAAAGTATTTTCGAGCGGAGG	180	
Qy	181	CTATGATGTTTTCGAGGCGACAGATGCGGCGGAAATGCATCAGATCCTCTCTGAATATGA	240	
Db	181	CTATGATGTTTTCGAGGCGACAGATGCGGCGGAAATGCATCAGATCCTCTCTGAATATGA	240	
Qy	241	CATCAACTGTTGATGATGATATCAATCTGCGCGGTGAAGACGGTCTTCTGTTAGCGCG	300	
Db	241	CATCAACTGTTGATGATGATATCAATCTGCGCGGTGAAGACGGTCTTCTGTTAGCGCG	300	
Qy	301	TGAATCGCGGAGCGAGCGAATGTTGCTGATGTTCTGATGCTGCGCGGTGACACGAGT	360	
Db	301	TGAATCGCGGAGCGAGCGAATGTTGCTGATGTTCTGATGCTGCGCGGTGACACGAGT	360	
Qy	361	CGATAAAATTTCTCGGCTCGAAATCGGTGCGATGATGATCACTACCAACCGTTCAACC	420	
Db	361	CGATAAAATTTCTCGGCTCGAAATCGGTGCGATGATGATCACTACCAACCGTTCAACC	420	
Qy	421	GGGTGAATGACGATTCGTGACGCAACCTACTGTCCTGATCCATGAAATCTGGGTACTGT	480	
Db	421	GGGTGAATGACGATTCGTGACGCAACCTACTGTCCTGATCCATGAAATCTGGGTACTGT	480	
Qy	481	CAGCGAAGACGTCGTAGCGTTGAAGCTTACAGTTCAATGTTGGGAACCTGGACATCAA	540	
Db	481	CAGCGAAGACGTCGTAGCGTTGAAGCTTACAGTTCAATGTTGGGAACCTGGACATCAA	540	
Qy	541	CAGCGGTTCTGTTGATCGGCGCTGATGGGAGCAGTACAAGTTCGCGCGCAGCGAGTTCG	600	
Db	541	CAGCGGTTCTGTTGATCGGCGCTGATGGGAGCAGTACAAGTTCGCGCGCAGCGAGTTCG	600	
Qy	601	CGCCATGCTTCACTTCTGTTGAAAAACCCAGGCAAAATTCAGTCCCGTGTGAACTGCTGAA	660	
Db	601	CGCCATGCTTCACTTCTGTTGAAAAACCCAGGCAAAATTCAGTCCCGTGTGAACTGCTGAA	660	
Qy	661	GAATATGACCGCGGTGAGTGAACCGCAGCAGTACTGAGACGTGACGATCCGCGCG	720	
Db	661	GAATATGACCGCGGTGAGTGAACCGCAGCAGTACTGAGACGTGACGATCCGCGCG	720	
Qy	721	TATTCGTAACATTTTCGAATCTACGCGCGATACGCGGAAATTCATCGCCACCATTCACGG	780	
Db	721	TATTCGTAACATTTTCGAATCTACGCGCGATACGCGGAAATTCATCGCCACCATTCACGG	780	
Qy	781	TGAAGGTTATCGCTTCTCGGCTGATCTGGAAGATTAATCGGCTTTACGACCGTCAAAAA	840	
Db	781	TGAAGGTTATCGCTTCTCGGCTGATCTGGAAGATTAATCGGCTTTACGACCGTCAAAAA	840	

Qy	841	AACGGCGCTTTTAGCGCGCTTTTATTTTCAACCTTATTTCCAGATACGTAACATCATC	900
Db	841	AACGGCGCTTTTAGCGCGCTTTTATTTTCAACCTTATTTCCAGATACGTAACATCATC	900
Qy	901	GTCCGTTGTAACTTCTTTTACTGGCTTTT	927
Db	901	GTCCGTTGTAACTTCTTTTACTGGCTTTT	927

RESULT 2

ABL90421/c

ID ABL90421 standard; cDNA; 1732 BP.

XX

AC ABL90421;

XX

DT 24-MAY-2002 (first entry)

XX

Human polynucleotide SEQ ID NO 983.

XX

Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
neurological disease; infection; human; secreted protein; gene; ss.

OS Homo sapiens.

XX

PN WO200190304-A2.

XX

PD 29-NOV-2001.

XX

PF 18-MAY-2001; 2001WO-US016450.

XX

PR 19-MAY-2000; 2000US-0205515P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Birse CE, Rosen CA;

XX

DR WPI; 2002-122018/16.

XX

DR P-PSDB; ABB90012.

XX

Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
prevention of neural, immune system, muscular, reproductive,
gastrointestinal, pulmonary, cardiovascular, renal and proliferative
disorders.

Claim 4; SEQ ID NO 983; 2081pp + Sequence Listing; English.

XX

The invention relates to novel genes (ABL89449-ABL90853) and proteins
(ABB89040-ABB90444) useful for preventing, treating or ameliorating
medical conditions e.g. by protein or gene therapy. The genes are
isolated from a range of human tissues disclosed in the specification.
The nucleic acids, proteins, antibodies and (ant)agonists are useful in
the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
ovarian cancer and other cancers of the adrenal gland, bone marrow,
breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
disorders e.g. Addison's disease, allergies, autoimmune haemolytic
anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
cardiovascular disorders such as myocardial ischaemia; (d) wound healing
; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
infectious diseases such as viral, bacterial, fungal and parasitic
infections. Note: The sequence data for this patent did not form part of
the printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 1732 BP; 439 A; 391 C; 417 G; 477 T; 0 U; 8 Other;

Query Match 99.9%; Score 925.8; DB 6; Length 1732;

Best Local Similarity 99.7%; Pred. No. 4.9e-29;

Matches 924; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GTCTGTTTACGCCGATCATGTTTAAATTTGTCAGCATCATCAGCAGGTGAGGACTTTTGT 60
Db |||||
1313 GTCTGTTTACGCCGATCATGTTTAAATTTGTCAGCATCATCAGCAGGTGAGGACTTTTGT 1254
QY 61 ACTTCTGTTTTCGATTTAGTTGGCAATTTAGTAGCAACATGACAGCCCGCATTTCT 120
Db |||||
1253 ACTTCTGTTTTCGATTTAGTTGGCAATTTAGTAGCAACATGACAGCCCGCATTTCT 1194
QY 121 TATCGTTGAAGACGAGTTGGTAAACGCAACACGTTTGAAGTATTTTGAAGCGGAAG 180
Db |||||
1193 TATCGTTGAAGACGAGTTGGTAAACGCAACACGTTTGAAGTATTTTGAAGCGGAAG 1134
QY 181 CTATGATGTTTTCGAAGCAGACAGATGGCGCGGAATGCAATCAGATCTCTCTGAATATGA 240
Db |||||
1133 CTATGATGTTTTCGAAGCAGACAGATGGCGCGGAATGCAATCAGATCTCTCTGAATATGA 1074
QY 241 CATCAACCTGGTGATCATGGATATCAATCTGCGCGGTGAAGACGCTCTCTGTAGCGG 300
Db |||||
1073 CATCAACCTGGTGATCATGGATATCAATCTGCGCGGTGAAGACGCTCTCTGTAGCGG 1014
QY 301 TGAACCTGCGCAGCAGCGCAATGTTGCGTTGATGTTCTCTGACTGGCGCGTGACAAAGT 360
Db |||||
1013 TGMWMTGCGCAGCAGCGCAATGTTGCGTTGATGTTCTGACTGGCGCGTGACAAAGT 954
QY 361 CGATAAATTTCTCGGCTCGAAATCGGTGACATGACTACATCAACAAACCGTTCAACC 420
Db |||||
953 CGATAAATTTCTCGGCTCGAAATCGGTGACATGACTACATCAACAAACCGTTCAACC 894
QY 421 GGGTGAACGATGCTGTCGACGCAACCTACTGTCGGTACCATGATCTGGTACTGTT 480
Db |||||
893 GGGTGAACGATGCTGTCGACGCAACCTACTGTCGGTACCATGATCTGGTACTGTT 834
QY 481 CAGCGAAGACGCTGCTAGCGTTGAAAGCTCAAAAGTCAATGTTGGGAACCTGCAACAA 540
Db |||||
833 CAGCGAAGACGCTGCTAGCGTTGAAAGCTCAAAAGTCAATGTTGGGAACCTGCAACAA 774
QY 541 CAGCGCTGCTGTTGATCGGCCCTGATGGCGAGCAGTACAAAGCTGCGCGCAGCAGTTCCG 600
Db |||||
773 CAGCGCTGCTGTTGATCGGCCCTGATGGCGAGCAGTACAAAGCTGCGCGCAGCAGTTCCG 714
QY 601 CGCCATGCTTCACTTCTGTGAAACCCAGGCAAAATTCAGTCCCGTCTGAACTCTGNA 660
Db |||||
713 CGCCATGCTTCACTTCTGTGAAACCCAGGCAAAATTCAGTCCCGTCTGAACTCTGNA 654
QY 661 GAAATGACCGCGGTGAGCTGAAACCGCAGCAGCTACTGTAGACGTGACGATCCGCG 720
Db |||||
653 GAAATGACCGCGGTGAGCTGAAACCGCAGCAGCTACTGTAGACGTGACGATCCGCG 594
QY 721 TATTCGTAAACATTTTGAATCTACGCGGATACGCGGAAATCATCGCCACCATTCACGG 780
Db |||||
593 TATTCGTAAACATTTTGAATCTACGCGGATACGCGGAAATCATCGCCACCATTCACGG 534
QY 781 TGAAGGTTATCGTTCTCGGTGATCTGGAAGATTAATCGGCTTTTACCAAGTCAAAAA 840
Db |||||
533 TGAAGGTTATCGTTCTCGGTGATCTGGAAGATTAATCGGCTTTTACCAAGTCAAAAA 474
QY 841 AACGGCGCTTTTATAGCGCGTTTATTTTATTTTCAACCTATTTTCCAGATACGTAATCTCATC 900
Db |||||
473 AACGGCGCTTTTATAGCGCGTTTATTTTATTTTCAACCTATTTTCCAGATACGTAATCTCATC 414
QY 901 GTCCGTTGTAATCTTTTACTGCTTT 927
Db |||||
413 GTCCGTTGTAATCTTTTACTGCTTT 387
```

RESULT 3
ACH96017

XX ACH96017 standard; DNA; 726 BP.

AC ACH96017;

XX 29-JUL-2004 (first entry)

XX XX

DE Klebsiella pneumoniae polynucleotide seqid 1812.
XX Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.
XX Klebsiella pneumoniae.
XX US6610836-B1.
XX 26-AUG-2003.
XX 27-JAN-2000; 2000US-00489039.
XX 29-JAN-1999; 99US-0117747P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Breton GL, Osborne M;
PI WPI; 2003-895346/82.
XX P-PSDB; ABO62466.
PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
PS Disclosure; SEQ ID NO 1812; 932pp; English.
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC pneumoniae polypeptide of the invention
XX
SQ . Sequence 726 BP; 189 A; 198 C; 189 G; 150 T; 0 U; 0 Other;

Query Match 62.3%; Score 577.2; DB 11; Length 726;
Best Local Similarity 87.2%; Pred. No. 4.6e-181;
Matches 633; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

```
QY 92 GTAGCAACATGACAGCCCGCATTCTTATCGTTGAAGACGAGTTGGTAAACAGCAAC 151
Db 1 GTAGCAACATGACAGCCCGCATTCTTATCGTTGAAGACGAGTTGGTAAACAGCAAC 60
QY 152 ACGTTGAAAGTATTTTCGAAGCGAGCTATGATGTTTTCGAAGCGACAGATGCGCG 211
Db 61 ACGTTGAAAGTATTTTCGAAGCGAGCTATGATGTTTTCGAAGCGACCGATGCGCG 120
QY 212 GAAATGATCAGATCCTCTCTGAATATGACATCAACCTGGTGATCATGATATCAATCTG 271
Db 121 GAAATGATCAGATCCTCTCTGAATATGATATCAACCTGGTGATCATGATATCAACCTG 180
QY 272 CCGGTAAGAACGCTTCTGTTAGCGGTGAACTGCGGAGCAGCGGATGTTGGTGTG 331
Db 181 CCGGTAAGAAATGCTTCTGCTGCGCGCGAGCTGCGGAGCAGCTGACGTCGCGCTG 240
QY 332 ATGTTCTCTGACTGCGCGTGACAAAGTTCGATAAAATTTCTCGGCTCGAAATCGGTGCA 391
Db 241 ATGTTCTTAAACGCGCGGACAAACGAGTGGATAAGATCTTGGCTCGAAATCGCGCT 300
QY 392 GATGACTATACCAAAACCGTTTCAACCGCGGTGAACTGACGATTCGTGACGCAACCTTA 451
Db 301 GACGACTATATCACTAAACCGTTTAAACCGCGGTGAACTGACTATCCGCGCGCAACCTG 360
QY 452 CTGTCCCGTACCATGAACTCTGGTACTGTACGCGAAGAACGTCGTAGCGTTGAAAGCTAC 511
Db 361 CTCTCCCGTACCATGAACTCTGGTACTGTACGCGAAGAACGTCGTAGCGTTGAAAGCTAC 420
QY 512 AAGTTCAATGTTGGAACTGGACATCAACAGCGGTTGTTTGTATCGGCCCTCATGCGCAG 571
Db 421 AAGTTCAACGCTGGGAACTCGATATCAACAGCCGTTCTCTGTTAGCCCGCAACCGGAA 480
```

QY	572	CAGTACAAAGCTCCGCGCAGCGAGTTTCGCGCCATGCTTCACTTCTGTGAAAACCCAGGC	631
Db	481	CAGTACAAAGCTCCGCGCAGCGAGTTTCGCGCCATGCTTCACTTCTGTGAAAACCCAGGC	540
QY	632	AAAATTCAAGTCCCGTCTGAACTGCTGAAGAAAATGACCGCGGTGAGCTGAAACCGCAC	691
Db	541	AAAATTCAAGTCCCGTCTGAACTGCTGAAGAAAATGACCGCGGTGAGCTGAAACCGCAT	600
QY	692	GACCGTACTGTAGAGTGAAGTACCGCGTATTCGTAAACATTTTCAATCTACGCGGAT	751
Db	601	GACCGTACTGTAGAGTGAAGTACCGCGTATTCGTAAACATTTTCAATCTACGCGGAT	660
QY	752	ACGCGCGAAATCATCCGACCAATTCACCGTGAAGGTTATCGCTTCTGCGGTGATCTGGAA	811
Db	661	ACGCGCGAAATCATCCGACCAATTCACCGTGAAGGTTATCGCTTCTGCGGTGATCTGGAG	720
QY	812	GATTAA	817
Db	721	GAATAA	726
RESULT 4			
ACF67367_28			
Continuation (29 of 57) of ACF67367 from base 2800001 (Photorhabdus luminescens nucleotide sequence split into 57 fragments) LOCUS ACF67367 Accession ACF67367			
WP	Fragment Name	Begin	End
WP	ACF67367_00	100001	110000
WP	ACF67367_01	100001	210000
WP	ACF67367_02	200001	310000
WP	ACF67367_03	300001	410000
WP	ACF67367_04	400001	510000
WP	ACF67367_05	500001	610000
WP	ACF67367_06	600001	710000
WP	ACF67367_07	700001	810000
WP	ACF67367_08	800001	910000
WP	ACF67367_09	900001	1010000
WP	ACF67367_10	1000001	1110000
WP	ACF67367_11	1100001	1210000
WP	ACF67367_12	1200001	1310000
WP	ACF67367_13	1300001	1410000
WP	ACF67367_14	1400001	1510000
WP	ACF67367_15	1500001	1610000
WP	ACF67367_16	1600001	1710000
WP	ACF67367_17	1700001	1810000
WP	ACF67367_18	1800001	1910000
WP	ACF67367_19	1900001	2010000
WP	ACF67367_20	2000001	2110000
WP	ACF67367_21	2100001	2210000
WP	ACF67367_22	2200001	2310000
WP	ACF67367_23	2300001	2410000
WP	ACF67367_24	2400001	2510000
WP	ACF67367_25	2500001	2610000
WP	ACF67367_26	2600001	2710000
WP	ACF67367_27	2700001	2810000
WP	ACF67367_28	2800001	2910000
WP	ACF67367_29	2900001	3010000
WP	ACF67367_30	3000001	3110000
WP	ACF67367_31	3100001	3210000
WP	ACF67367_32	3200001	3310000
WP	ACF67367_33	3300001	3410000
WP	ACF67367_34	3400001	3510000
WP	ACF67367_35	3500001	3610000
WP	ACF67367_36	3600001	3710000
WP	ACF67367_37	3700001	3810000
WP	ACF67367_38	3800001	3910000
WP	ACF67367_39	3900001	4010000
WP	ACF67367_40	4000001	4110000
WP	ACF67367_41	4100001	4210000
WP	ACF67367_42	4200001	4310000
WP	ACF67367_43	4300001	4410000
WP	ACF67367_44	4400001	4510000
WP	ACF67367_45	4500001	4610000
WP	ACF67367_46	4600001	4710000

WP	ACF67367_47	4700001	4810000
WP	ACF67367_48	4800001	4910000
WP	ACF67367_49	4900001	5010000
WP	ACF67367_50	5000001	5110000
WP	ACF67367_51	5100001	5210000
WP	ACF67367_52	5200001	5310000
WP	ACF67367_53	5300001	5410000
WP	ACF67367_54	5400001	5510000
WP	ACF67367_55	5500001	5610000
WP	ACF67367_56	5600001	5648894
Query Match			
55.1%; Score 510.6; DB 10; Length 110000;			
Best Local Similarity 77.5%; Pred. No. 1.3e-157;			
Matches 618; Conservative 0; Mismatches 179; Indels 0; Gaps 0;			
QY	44	AGGTGAGGAGCTTTGTACTTCTCTGTTTTCGATTTTGTGGCAATTTAGGTAGCAAAATG	103
Db	67943	AGATAAGAGTAGTTTGAATCAATCTCGATTTATTTTAAAGCGATTTCTAGGTAGCAACCATG	68002
QY	104	CAGACCCCGCACATTTCTATCGTTTGAGACAGTTGGTAACACGCAACACGTTGAAAAGT	163
Db	68003	CAACCCCGCACATTTTGAATTTTGAGACGAAATTTGTCATCTCGCAATACCTTAANAAGC	68062
QY	164	ATTTTCGAAGCGGAGGCTATGATGTTTTCGAAGCGACAGATGGCGGAAATGTCATCAG	223
Db	68063	ATTTTCGAAGCTGAAGGGTACATAGTTTATGAAGCCACTGATGGTTTCAGAAATGCACCAT	68122
QY	224	ATCTCTCTGAATATGACATCAACTGATGATCATGATATCAATCTCCGGGTGAAGAAC	283
Db	68123	ATTTCTGTCAAAACAGACATTAACCTGGTATTTATGGAATTAATTAACCTTCCAGGCAAAAT	68182
QY	284	GGTCTTCTGTAGCGCGTGAAGTCTGCGGAGCAGGCGAATGTTGCGTGTGATGTTCTCTGACT	343
Db	68183	GGTCTTCTGTAGCGCGGAACTGGAGAGCAAGCAATGTTGCTTGTGATGTTCTCTGACT	68242
QY	344	GGCGGTGACAAACGAGTGCATAAAATTTCTCGGCCCTCGAAATCGGTGTCAGATGACTACATC	403
Db	68243	GGCGGTGATACGAAGTAGACAAAATCTTGGCTTGGAAATCGGTGTCAGATGATTACATC	68302
QY	404	ACCAAAACCGTTCAACCCCGTGAAGTACGAGTATGTCGACGCAACCTACTGTCCGCTACC	463
Db	68303	ACCAAAACCAATTAATCCGCGTGAATTTGACTATTTGCGGCCCGCAACCTGCTTCTCGTACC	68362
QY	464	ATGAATCTGGGTACTGTGTCAGGAAAGCTGTCAGCGTTGAAAGCTACAAGTTCAATGCT	523
Db	68363	ATGAATCTGAGTAAATGTGTCAGGAAAGCTGCCAAGTTGAGAGTTATTAATTTCAACGCT	68422
QY	524	TGGGAACTGGACATCAACAGCCGTTTCGTTGATCGGCCCTGATGGCGAGCAGTACAAGCTG	583
Db	68423	TGGGAACTTAGATATCAATAGCCGTTCTCTGATAAGCCCGCTGGTGAGCCGTATAAATTG	68482
QY	584	CGCGCAGCGAGTTCCGCGCCATGCTTCACTTCTGTGAAAACCCAGGCAAAATTCAGTCC	643
Db	68483	CCAGCAGCGAGTTCCGCGCCATGCTTCACTTCTGTGAAAATCCAGGCAAAATTCACAAACA	68542
QY	644	CGTCTGAACTGTCGAAGAAAATGACCGCGCTGAGCTGAAACCCGACGACCTGCTGTA	703
Db	68543	CGCGCAGTTTACTGAAAAGATGACAGCTGCTGAATTTGAAACCTCAACGACCTACTGTT	68602
QY	704	GACGTGACGATCCGCGCTTATTCGTAACATTTTCGAATCTACGCCGATACGCCGGAATTC	763
Db	68603	GATGTAACGATTCGTCGTATCCGTAACATTTTGAATCGAGCCAGATACGATGGAATTT	68662
QY	764	ATCGCCACCATTCACGGTGAAGGTTATCGCTTCTCGGTGATCTCGAAGATTAATCGGCT	823
Db	68663	ATCGCCACCATTCACGGTGAAGGTTATCGCTTCTCGGTGATTTGGAAGCATAACTGAA	68722
QY	824	TTACCACCGTCAAAAAA	840
Db	68723	CTATCGGTCTCTAACAA	68739

ACF65386_0/c

WP Sequence split into 7 fragments LOCUS ACF65386 Accession Acf65386

WP	Fragment Name	Begin	End
WP	ACF65386_0	1	110000
WP	ACF65386_1	100001	210000
WP	ACF65386_2	200001	310000
WP	ACF65386_3	300001	410000
WP	ACF65386_4	400001	510000
WP	ACF65386_5	500001	610000
WP	ACF65386_6	600001	700779

ID ACF65386 standard; DNA; 700779 BP.

XX

AC ACF65386;

XX

XX 20-NOV-2003 (first entry)

XX

XX Photorhabdus luminescens nucleotide sequence #39.

XX

Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 detection; food; gene expression; plant; animal; microorganism; toxin;
 antibiotic; biopesticide; virulence factor; disease model; plague;
 whooping cough; gene; ds.

XX

XX Photorhabdus luminescens.

XX

XX WO200294867-A2.

XX

XX 28-NOV-2002.

XX

XX 07-FEB-2002; 2002WO-IB003040.

XX

XX 07-FEB-2001; 2001PR-00001659.

XX

XX (INSP) INST PASTEUR.

XX

XX (CNRS) CNRS CENT NAT RECH SCI.

XX

XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;

XX Buchrieser C;

XX WPI; 2003-148459/14.

XX

XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,

XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX

XX Claim 1; SEQ ID NO 39; 1205pp; French.

XX

XX The invention relates to the isolation of genes and their encoded

XX proteins from Photorhabdus luminescens. The isolated sequences are

XX sources of probes and primers for detecting the genome of P. luminescens

XX and related species; to study polymorphisms; for gene analysis and for

XX detection/amplification of the genes. Antibodies (Ab) raised against the

XX polypeptides encoded by the genes are used for detection/identification

XX of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that

XX carry a gene-containing vector are used to select compounds that

XX modulate, regulate, induce or inhibit expression of the genes in plants,

XX animals or microorganisms other than P. luminescens and are able to alter

XX response or sensitivity to toxins and antibiotics produced by P.

XX luminescens. Cells transformed to express the genes are useful for

XX recombinant production of the proteins, particularly toxins and

XX antibacterials useful as insecticides, bactericides and fungicides. The

XX genes, proteins, vectors containing the genes and Ab are also useful

XX therapeutically (to treat microbial infection by bacteria or fungi that

XX are sensitive to P. luminescens-encoded toxins or antibiotics) and as

XX biopesticides. Other uses of the genes and the proteins are as virulence

XX factors and for identifying targets of human diseases for which P.

XX luminescens is a model (particularly plague and whooping cough). This

XX sequence represents one of the isolated P. luminescens genes

XX

XX Sequence 700779 BP; 190440 A; 140977 C; 164444 G; 204917 T; 0 U; 1 Other;

XX

XX Query Match 55.1%; Score 510.6; DB 10; Length 110000;

XX Best Local Similarity 77.5%; Pred. No. 1.3e-157;

XX Matches 618; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY	44	AGGT	CAGG	CACTTT	TTG	TACT	TTCT	CTG	TTT	CG	CAAT	TTAG	TTGG	CAAT	TTAG	TAG	CAAA	CAATG	103
Db	30684	AGATA	AGAG	TAGT	TTT	GAAT	CAT	TTCT	CG	CAAT	TTAG	TTGG	CAAT	TTAG	TAG	CAAA <td>CAATG</td> <td>30625</td>	CAATG	30625	
QY	104	CAG	CCCCG	CACT	TTT	TAT	CGT	TTG	AA	GAC	AG	TTG	TTG	TA	CA	CG	CA	CACT	163
Db	30624	CAAA	CCCCG	CACT	TTT	TAT	CGT	TTG	AA	GAC	AG	TTG	TTG	TA	CA	CG	CA	CACT	30565
QY	164	ATTTT	CGA	AG	CGG	AT	ATG	ATG	TTT	CG	CA	AG	CGG	AT	ATG	ATG	TTT	CG	223
Db	30564	ATTTT	CGA	AG	CGG	AT	ATG	ATG	TTT	CG	CA	AG	CGG	AT	ATG	ATG	TTT	CG	30505
QY	224	ATCCT	CTCT	GAAT	ATG	ATC	CA	CTG	TTG	AT	ATG	ATG	TTT	CG	CA	AG	CGG	AT	283
Db	30504	ATTTT	CGA	AG	CGG	AT	ATG	ATG	TTT	CG	CA	AG	CGG	AT	ATG	ATG	TTT	CG	30445
QY	284	GGT	CTT	CTG	TAG	CGG	TTG	AA	GAC	AG	TTG	TTG	TA	CA	CG	CA	CACT	343	
Db	30444	GGT	CTT	CTG	TAG	CGG	TTG	AA	GAC	AG	TTG	TTG	TA	CA	CG	CA	CACT	30385	
QY	344	GGC	GTG	CA	AC	GA	AG	CGG	TTG	AA	GAC	AG	TTG	TA	CA	CG	CA	CACT	403
Db	30384	GGC	GTG	CA	AC	GA	AG	CGG	TTG	AA	GAC	AG	TTG	TA	CA	CG	CA	CACT	30325
QY	404	ACC	AA	CCG	TT	CA	AC	CGG	TTG	AA	GAC	AG	TTG	TA	CA	CG	CA	CACT	463
Db	30324	ACC	AA	CCG	TT	CA	AC	CGG	TTG	AA	GAC	AG	TTG	TA	CA	CG	CA	CACT	30265
QY	464	ATG	AAT	CTG	GT	AT	CA	CTG	TTG	AA	GAC	AG	TTG	TA	CA	CG	CA	CACT	523
Db	30264	ATG	AAT	CTG	GT	AT	CA	CTG	TTG	AA	GAC	AG	TTG	TA	CA	CG	CA	CACT	30205
QY	524	TGG	GA	CTG	GA	CA	TC	CA	CGG	TTG	AA	GAC	AG	TTG	TA	CA	CG	CA	583
Db	30204	TGG	GA	CTG	GA	CA	TC	CA	CGG	TTG	AA	GAC	AG	TTG	TA	CA	CG	CA	30145
QY	584	CCG	CG	CA	CGG	TTG	AA	GAC	AG	TTG	TA	CA	CGG	TTG	AA	GAC	AG	TTG	643
Db	30144	CCG	CG	CA	CGG	TTG	AA	GAC	AG	TTG	TA	CA	CGG	TTG	AA	GAC	AG	TTG	30085
QY	644	CGT	CTG	TA	CTG	TA	CA	CTG	TTG	AA	GAC	AG	TTG	TA	CA	CG	CA	CACT	703
Db	30084	CGC	CA	GA	TTT	TA	CTG	TA	CA	CTG	TTG	AA	GAC	AG	TTG	TA	CA	CACT	30025
QY	704	GAC	GTG	CA	CTG	TA	CTG	TA	CA	CTG	TTG	AA	GAC	AG	TTG	TA	CA	CACT	763
Db	30024	GAT	GTG	CA	CTG	TA	CTG	TA	CA	CTG	TTG	AA	GAC	AG	TTG	TA	CA	CACT	29965
QY	764	ATC	CC	CA	CTG	TA	CTG	TA	CA	CTG	TTG	AA	GAC	AG	TTG	TA	CA	CACT	823
Db	29964	ATC	CC	CA	CTG	TA	CTG	TA	CA	CTG	TTG	AA	GAC	AG	TTG	TA	CA	CACT	29905
QY	824	TTAC	CA	CGG	TTG	AA	GAC	AG	TTG	TA	CA	CGG	TTG	AA	GAC	AG	TTG	TA	840
Db	29904	CTAT	CGG	TTG	AA	GAC	AG	TTG	TA	CA	CGG	TTG	AA	GAC	AG	TTG	TA	CA	29888

RESULT 6

ADJ38390

ID ADJ38390 standard; DNA; 759 BP.

XX

XX AC ADJ38390;

XX

XX 06-MAY-2004 (first entry)

XX

XX Pantoea ananatis arca (aerobic respiration control) gene SeqID19.

XX

XX gamma-proteobacterium; Arca; aerobic respiration control;

XX Arca protein function; chromosomal arca gene; fermentation; L-amino acid;

XX L-lysine; L-glutamic acid; gene; ds.

XX

XX Pantoea ananatis.

XX

```

XX Key Location/Qualifiers
FH CDS 41..757
FT /*tag= a
FT /product= "pantoea ananatis arca protein"
XX
PN BP1382686-A1.
XX
XX 21-JAN-2004.
XX
XX 11-JUL-2003; 2003EP-00015911.
XX
XX 12-JUL-2002; 2002JP-00203764.
XX
XX (AJIN ) AJINOMOTO CO INC.
XX
XX Ishikawa Y, Imaizumi A, Matsui K, Kojima H;
XX WPI; 2004-192975/19.
XX P-PSDB; ADJ38391.
XX
XX New bacterium useful for producing a target substance (especially L-
XX lysine, L-glutamic acid or another L-amino acid) is modified so that the
XX Arca protein does not normally function.
XX
XX Example 4; SEQ ID NO 19; 38pp; English.
XX
XX This invention relates to a novel gamma-proteobacterium able to produce a
XX target substance such as an amino acid which is modified so that an Arca
XX (aerobic respiration control) protein does not normally function. The
XX disruption to Arca protein function is preferably due to disruption of a
XX chromosomal arca gene. The bacterium and method are useful in producing a
XX target substance by fermentation. The target substance is preferably an L
XX -amino acid, especially L-lysine or L-glutamic acid. Production of the
XX target substance using the modified bacterium is more efficient. The
XX present sequence is that of the P ananatis gene which encodes the Arca
XX protein and which was used in the exemplification of the invention.
XX
XX Sequence 759 BP; 205 A; 192 C; 183 G; 179 T; 0 U; 0 Other;
XX
XX Query Match 54.9%; Score 508.6; DB 12; Length 759;
XX Best Local Similarity 80.6%; Pred. No. 3.5e-158;
XX Matches 607; Conservative 0; Mismatches 144; Indels 2; Gaps 1;
XX
XX 69 TTTTCATTAGTTGGC--AATTTAGTAGCAACATGACAGCCCGCACATTTCTTATCGT 126
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 7 TTTTCATTAGTTGGCAAAATTTAGTAGTAACATGACAGCCCGCACATTTCTATCGT 66
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 127 TGAAGACGAGTTGGTAACACGCAACACGTTGAAAAGTATTTTCGAAGCGGAGGCTATGA 186
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 67 TGAAGACGAACTGGTCACGCGCAATACCTCAAAAGCATTTTGGAGCGGAAGGTTATGT 126
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 187 TGTGTTTCGAAGCAGATGCGCGGAATATGATCAGATCTCTCTGTAATATGATCA 246
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 127 CGTGACGAAGCGACCGATGGTGCAGAGATGCACGAGGTGTTGACCGCAATATGATCAA 186
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 247 CCTGTGTATCATGGATATCAATCTCGGGTAAGAACGCTCTTCTGTAGCGCGTA 306
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 187 TCTGGTTATTTAGGACATCAATCTCGCGGGTAAACCGGCCCTGTTACTGGACGTA 246
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 307 GCGGAGCAGCGGAATGTTGCGTTGATGTTCTCTGACTGGCCCGTGACACGAGTGCATAA 366
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 247 GCGTGAGCAAGCAATGTGCGATTGATGTTCTGACCGGACGCGATACGGAATGCATAA 306
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 367 AATTCGCGCTCGAAATCGGTGCGAGATGACATACCAACCGTTCAACCCCGTGA 426
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 307 AATTCCTGGCGTGGAAATTTGGTGCAGACGACTACATTAAGAGCCGTTTAACCCACGGA 366
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 427 ACTGACGATTCGTGCACGCAACCTACTGTCGCGTACCAGTATCTGGGTACTGTGACGGA 486
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 367 ATTAATAATTCGTGCACGTAACCTGCTGTTGGCGACCAATGAATTTGCCCTTACCCCAATGA 426
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 487 AGAACGTGCTAGCGTTGAAAGCTACAAGTTCAATGTTGGGAACTGGGAACTCAACACGCGG 546
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 7

ACF69844

ID ACF69844 standard; DNA; 717 BP.

XX ACF69844;

AC ACF69844;

XX ACF69844;

DT 20-NOV-2003 (first entry)

XX 20-NOV-2003 (first entry)

DE Photorhabdus luminescens nucleotide sequence #8311.

XX Photorhabdus luminescens nucleotide sequence #8311.

KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;

KW detection; food; gene expression; plant; animal; microorganism; toxin;

KW antibiotic; biopesticide; virulence factor; disease model; plague;

KW whooping cough; gene; ds.

XX whooping cough; gene; ds.

XX Photorhabdus luminescens.

XX Photorhabdus luminescens.

XX WO200294867-A2.

XX 28-NOV-2002.

XX 07-FEB-2002; 2002WO-IB003040.

XX 07-FEB-2001; 2001FR-00001659.

XX (INSP) INST PASTEUR.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;

XX Buchrieser C;

XX WPI; 2003-148459/14.

XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,

XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX Claim 2; SEQ ID NO 8311; 1205pp; French.

XX The invention relates to the isolation of genes and their encoded

XX proteins from Photorhabdus luminescens. The isolated sequences are

XX sources of probes and primers for detecting the genome of P. luminescens

XX and related species; to study polymorphisms; for gene analysis and for

XX detection/amplification of the genes. Antibodies (Ab) raised against the

XX polypeptides encoded by the genes are used for detection/identification

XX of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that

XX carry a gene-containing vector are used to select compounds that

XX modulate, regulate, induce or inhibit expression of the genes in plants,

XX animals or microorganisms other than P. luminescens and are able to alter

XX response or sensitivity to toxins and antibiotics produced by P.

CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens genes
XX
XX Sequence 717 BP; 213 A; 158 C; 158 G; 188 T; 0 U; 0 Other;

Query Match 53.2%; Score 493; DB 10; Length 717;
Best Local Similarity 80.5%; Pred. No. 5.4e-153;
Matches 577; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 101 ATGCAAGCCCGCACATCTTATCGTTGAAGCAGATGGTGAACACGCAACAGTGTGAAA 160
Db 1 ATGCAAGCCCGCACATCTTATCGTTGAAGCAGAAATGTCTCACTCGCAATACCCCTAAAA 60
QY 161 AGTATTTTCGAAGCGGAAGGCTATGATGTTTTCGAAGCGACAGATGGCGCGAATGCAT 220
Db 61 AGCATTTTCGAAGCTGAAGGCTACATGTTTATGAAGCACTGATGGTTCAAGAAATGCAC 120
QY 221 CAGATCTCTCTGAAATATGACATCAACCTGTGTGATCATGGATATCAATCTGCCGGTAA 280
Db 121 CATATCTGTCAACACACGACATTAACCTGTGTGATGATTAATTAACCTCCAGGCAAA 180
QY 281 AACGCTCTCTGTAGCGGTGAACCTGCGGAGCAGCGGAATGTTGCGTTGATGTTCTGT 340
Db 181 AATGCTTGTGTACTTTGCCCGCAACTGCGGAGCAAGCAAAATGTTGCTTTGATGTTCTGT 240
QY 341 ACTGCCGTGACCAAGAGTGCATATAAATCTCGSCCTCGAAATCGGTGCGAGATCACTAC 400
Db 241 ACTGCCGTGATACGAAGTACAAAATCTTGGCTTGGAAATCGGTGCGAGATGATTAC 300
QY 401 ATCACCACACCGTTCAACCGCGTGAACCTGACGATTCGTGACGCAACCTACTGTCCGT 460
Db 301 ATCACCACACCAATTAATCCGCGTGAATGACTATTCGCGCCGCAACCTGTTCTGT 360
QY 461 ACCATGATCTGGTACTGTGACGAAGAACGTCTAGCGTTGAAAGCTACAGTTCAAT 520
Db 361 ACCATGATCTGAGTAATGTACGCGAGGAACGTCTGCCAAGTTGAGAGTTATAAATTCAC 420
QY 521 GTTTCGGAACCTGACATCAACAGCGCTTCGTTGATCGGCCCTGTATGGCGAGCAGTACAG 580
Db 421 GTTTCGGAATTGATATCAATAGCCGTTCTCTGATAAGCCCGCTGGTGAGCCGTATAAA 480
QY 581 CTGCGCGCAGCGAGTTCCGCGCCATGCTTCACTTCTGTGAAAAACCCAGGCAAAATTCAG 640
Db 481 TTGCCACGCGAGTTCCGCGCCATGCTTCACTTCTGTGAAAAATCCAGGCAAAATTCAG 540
QY 641 TCCCGTGTGAACCTGCTGAAGAAATGACCGCGCTGAGCTGAAACCGCGACGACGCTACT 700
Db 541 ACACGCGCAGATTACTGAAAAAGATGACAGGCTCGTGAATGAAACCTCAGCAGCGTACT 600
QY 701 GTGACGCTGACGATCCGCGTATTTCGTAAACATTTTCAATCTACGCGCGATACCGCGAA 760
Db 601 GTTGATGTAACGATTCGTCGTATCCGTAAACATTTTGAATGACGCCAGATACGATGAA 660
QY 761 ATATCGGCCACCATTCAGCGTGAAGGTTATCGCTTCTCGCGGTGATCTCGGAAGATTA 817
Db 661 ATTATCGGCCACCATTCAGCGTGAAGGCTATCGTTTCTGTGTTGATTTGGAAGCATAA 717

RESULT 8

ADFO1748

ID ADFO1748 standard; DNA; 726 BP.

XX

AC ADFO1748;

XX

DT 12-FEB-2004 (first entry)

XX Bacterial polynucleotide #2033.
DE
XX
XX Proteus mirabilis infection; bacterial infection; antibacterial;
KW immunostimulant; gene; ds.
XX
XX Proteus mirabilis.
OS
XX
XX US6605709-B1.
PN
XX
XX 12-AUG-2003.
PD
XX
XX 05-APR-2000; 2000US-00543681.
PF
XX
XX 09-APR-1999; 99US-0128706P.
PR
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Breton GL;
PI
XX
XX WPI; 2003-895291/82.
DR
XX
XX P-PSDB; ADF05920.
DR
XX
XX New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.
PT
XX
XX Disclosure; SEQ ID NO 2033; 870pp; English.
PS
XX
XX The invention relates to new Proteus mirabilis polypeptides and
CC polynucleotides. The invention also relates to antibodies against the
CC polypeptides, methods for producing the polypeptides, a method of
CC generating vaccines for immunising an individual against P. mirabilis, a
CC method for evaluating a compound for the ability to bind a P. mirabilis
CC polypeptide and a method for screening test compounds for anti-bacterial
CC activity. The polypeptides and polynucleotides are useful as molecular
CC targets for diagnosing, preventing and treating pathological conditions
CC resulting from bacterial infection, as reagents for diagnosis of
CC bacterial diseases, as components of antibacterial vaccines, as targets
CC for antibacterial drugs or as bio-control agents for plants. This
CC sequence represents a Proteus mirabilis polynucleotide of the invention.
XX
SQ Sequence 726 BP; 224 A; 137 C; 149 G; 216 T; 0 U; 0 Other;

Query Match 48.6%; Score 450.8; DB 10; Length 726;
Best Local Similarity 76.3%; Pred. No. 6.3e-139;
Matches 554; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

QY 92 GTAGCAACATGACAGACCCCGCACATCTTATCGTTGAAGCAGATTCGTTAACAGCAAC 151
Db 1 GTAGCAATATGCAAAACCCCGCACATCTGATTTGTAAGATGAAGTAGTTACTCGTAAT 60
QY 152 ACGTGTGAAAAGTATTTTGAAGCGGAAGGCTATGATGTTTTCGAAGCGACAGATGGCGG 211
Db 61 ACCCTGAAAAGCATATTTCGAAGCTGAAGGGTATATCGTACACGAAGCCACTGATGGCAAC 120
QY 212 GAAATGATCAGATCCCTCTGTAATATGACATCAACCTGCTGATCATCATGATATCAATCTG 271
Db 121 GAGATGCAATAATTTCTGTCCGACCATGATATCAATCTGGTCATTATGATTAATCTT 180
QY 272 CCGGTTAAGAACGGTCTTCTGTAGCGGTGAACTGCGCGAGCAGCGGCAATGTTTCGGTTG 331
Db 181 CCTGTTAAAAATGGTCTTCTATTAGCCGCTGAATTCGTGAACAGGTGAAGTGTTCATTA 240
QY 332 ATGTTCTTGACTGGCCGTGACCAAGAGTGCATATAAATTTCTCGGCTCGAATTCGGTGA 391
Db 241 ATGTTCTTAACAGGTCTGTGATAATGAAGTTCGATAAAATCTTAGGCTTTGAAATTCGGTCC 300
QY 392 GATGACTACATCACCAACCGTTCAACCGGCTGACTGACGATTCGTCGACGCAACCTTA 451
Db 301 GATGATTACATCACTAAACCAATTTAATCTCTGTTGAATTAATCTCCGTCGTAACCTTA 360

PF 05-MAR-2004; 2004WO-US007001.

XX 06-MAR-2003; 2003US-0453134P.

XX (CHIL-) CHILDRENS HOSPITAL INC.

XX Bakaletz LO, Munson RS, Dyer DW;

XX WPI; 2004-662422/64.

XX New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT useful for treating or preventing NTHi bacterial infections of the middle
PT ear and/or nasopharynx.

XX Claim 1; SEQ ID NO 683; 88bp; English.

XX The invention comprises nucleotide sequences (genes) from the genome of a
CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC sequences of the invention are useful for treating or preventing NTHi
CC bacterial infections of the middle ear and/or nasopharynx. The present
CC nucleic acid represents an NTHi DNA sequence of the invention.

XX SQ Sequence 19196 BP; 59302 A; 35787 C; 37096 G; 59811 T; 0 U; 0 Other;

Query Match 34.6%; Score 320.6; DB 13; Length 191996;

Best Local Similarity 66.0%; Pred. No. 3.7e-94;

Matches 480; Conservative 0; Mismatches 244; Indels 3; Gaps 1;

QY 85 AATTAGGTAGCAATGCAGACCCCGCACATTTCTTATCGTTGGAAGCAGTTCGTAAAC 144

Db 55869 AGTAAAGATCGAAATGACTCTCCAAAATTTCTCGTTGTTGAAGATGAAATTTGTCAC 55810

QY 145 ACGCAACACGTTGAAAGATTTTCGAAGCGAAGCTATGATGTTTTCGAAGCGACAGA 204

Db 55809 TCGAAATACGCTTAAGGATTTTGAAGCGAAGATATGATGTTTGAAGCGAAGAAA 55750

QY 205 TGGCGCGGAAATGCATCAGATCCTCTCTGAATATGACATCAACCTGGTGATCATCGATAT 264

Db 55749 TGGTGTGAAATGCATATATATGGCAATCATAATTAATTAATTTGGTTGATGGATAT 55690

QY 265 CAATCTGCGGTAGAACGGTCTTCTGTAGCGGTGAACCTGCGGAGCAGCGCAATGT 324

Db 55689 TAAATTCACGCAAAAACGGCTTATTAATGGCAAGAACCTCCGTGAAGAAATTAAGCTT 55630

QY 325 TCGCTTGATGTTCTGACTGGCGGTGACAAAGAGTCGATAAAATTCGGGCTCGAAAT 384

Db 55629 ACCTCTATTTTTTAACTGCTGCGAGATATGAGTGGATAAAATTTAGGTCTGAAAT 55570

QY 385 CGGTGCAGATGACTACATCAACAAACCGTTCAACCCGCGTGAACTGACGATTCGTGCACG 444

Db 55569 TGGTCCGACGATTTATTAACCAACCTTTTAAACCTAGAGAACTGACTATCCGCGCAG 55510

QY 445 CAACTACTGTCCTGACATGAATCTGGGTACTGTGACGAGAACGTCGTAGGTTGA 504

Db 55509 TAACTATATGCAATCGCAATG---TCGCATCAAGAAAAGAAATATACATTTGTCGAGA 55453

QY 505 AAGCTACAAGTTCAATGGTTGGGAATCGACATCAACAGCCGTTGTTGATCGGCGCTGA 564

Db 55452 ATTCTATCGTTTAAATGGTTGAAATTTAGACCTAAATAGCCACAGTTTAAATACACAGA 55393

QY 565 TGGCGAGCAGTACAAGCTGCCGCGAGCGAGTTCCGCGCCATGCTTCACTTCTGTGAAAA 624

Db 55392 AGGACAAGAAATTCAAACTTCTCGCAGTGAATTTGTCGAATGTTTACATTTCTGTGAAA 55333

QY 625 CCCAGGCAAAATTCAGTCCGCTGCTGAATCTGCTGAAGAAATGACCGCGCGTGAGCTGAA 684

Db 55332 TCCAGGAAATTTGCAACCGCGAAGAAATTTGCTGAAAAAATGACGGGACGAGGTTAAA 55273

QY 685 ACCGACACGCGTACTGTAGAGCTGACGATCCGCGTATTCGTAAACATTTGCAATCTAC 744

Db 55272 ACCTCAGATGTCACCGTAGATGTACAAATTCGAGTATCAGAAACATTTTGAAGATCA 55213

QY 745 GCCGATACGCGGAAATCATCGCCACCAATTCACGCGTGAAGGTTATCGTCTTCGCGTGA 804

Db 55212 TCCTAATCTCCAAATATCATTTATGACTATATCATGAGAGGCTATCGTTTTGCGGAGA 55153
QY 805 TCTGGAA 811
Db 55152 TATTGAA 55146

RESULT 11

AAT42063_09/c

Continuation (10 of 19) of AAT42063 from base 900001 (Haemophilus influenzae complete ge
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063

WP Fragment Name Begin End

WP AAT42063_00 1 110000

WP AAT42063_01 100001 210000

WP AAT42063_02 200001 310000

WP AAT42063_03 300001 410000

WP AAT42063_04 400001 510000

WP AAT42063_05 500001 610000

WP AAT42063_06 600001 710000

WP AAT42063_07 700001 810000

WP AAT42063_08 800001 910000

WP AAT42063_09 900001 1010000

WP AAT42063_10 1000001 1110000

WP AAT42063_11 1100001 1210000

WP AAT42063_12 1200001 1310000

WP AAT42063_13 1300001 1410000

WP AAT42063_14 1400001 1510000

WP AAT42063_15 1500001 1610000

WP AAT42063_16 1600001 1710000

WP AAT42063_17 1700001 1810000

WP AAT42063_18 1800001 1830121

Query Match 34.5%; Score 319.4; DB 2; Length 110000;

Best Local Similarity 65.8%; Pred. No. 6.8e-94;

Matches 480; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

QY 85 AATTAGGTAGCAACATGCAGACCCCGCACATTTCTTATCGTTGAAGCAGGTTGCTAAC 144

Db 36640 AGTATAGGATGGAATAATGACTACTCCAAAATTTCTCGTTGTTGAAGATGAATTTGTCAC 36581

QY 145 ACGCAACACGTTGAAAAAGTATTTTCGAAGCGAAGGCTATGATGTTTTCGAAGCGACAGA 204

Db 36580 TCGAAATACGCTTAAAGGATTTTTCGAAGCGAAGGATATGATGTTTTCGAAGCGAAGAAA 36521

QY 205 TGGCGGGAATGATCAGATCCTCTCTGAATATGACATCAACCTGGTGATCATCGATAT 264

Db 36520 TGGTGTGAAATGCAATCATATATTGGCAAAATCATAATATTAAATTTGGTTGTGATGATAT 36461

QY 265 CAATCTGCGGTTAAGAACGGTCTTCTGTAGCGGTGAACTGCGCGAGCAGGCGAAATGT 324

Db 36460 TAAATTCACGCAAAAACGGCTTATTAATTCGCAAGAGAACTCCGTGAAGAAATTAAGCTT 36401

QY 325 TCGCTTGATGTTCTGACTGCGCGGTGACAAAGTCGATAAAATTTCTCGGCTCGAAAT 384

Db 36400 ACCTCTATTTTTTAACTGCTGCGAGATAATGAAGTGGATATAATTTTAGGTTTGGAAAT 36341

QY 385 CGGTGCAGATGACTACATCACAAACCGTTCAACCCGCGTGAACCTGACGATTCGTGCGACG 444

Db 36340 TGGTGCAGATGACTATTTGACTAAAGCCCTTTTAAACCTAGAGAACTGACTATCCGCGCAG 36281

QY 445 CAACTACTGTCCTGACTACATGAATCTGGGTACTGTCTAGCGAAGAACGTCGTAGCGTTGA 504

Db 36280 TAACTATTGATGCTGCAATG---CCGCATCAAGAAAAGAAATACATTTTGGTCGAGA 36224

QY 505 AAGCTACAAGTTCAATGGTTGGGAATCTGGACATCAACAGCCGTTGTTGATCGGCCCTGA 564

Db 36223 ATTCTATCGTTTAAATGGTTGAAATTTAGACCTTAAATAGCCACAGTCTTAAATTTACACAGA 36164

QY 565 TGGCGAGCAGTACAAGCTGCGCGCAGGAGTTCGCGCCATGCTTCACTTCTGTGAAAA 624

Db 36163 AGGACAAGAAATTCAAACTTCTCGCAGTGAATTTCTCGCAATGTTTACATTTCTGTGAAAA 36104

QY 625 CCCAGCAAAATTTCAGTCCGGTGTGAATGCTGAAGAAATGACCGCCGGTGAAGTCAA 684
|||||
Db 36103 TCCAGAAATTTGCAACGCGGAAGATTTCTGAAAAAATGACCGACGAGAGTTAAA 36044
|||||
QY 685 ACCGCACGACCGTACTGTAGAGCGTGACGATCGCGGTATTCGTAAACATTTTGAATCTAC 744
|||||
Db 36043 ACCTCAGGATCGTACCGTAGATGTCAAAATTCGACGTATCAGAAACATTTTGAAGATCA 35984
|||||
QY 745 GCCGATACGCGGAAATCATCGCCACCATTCACGCTGAAGTTATCGTCTCGGGTGA 804
|||||
Db 35983 TCCCAATATCTCAAAATATCATTTATGACTATATACATGAGAGGCTATCGTTTTTGTGGAGA 35924
|||||
QY 805 TCTGGAAGA 813
|||||
Db 35923 TATTGAATA 35915
|||||

RESULT 12

AAS53409

ID AAS53409 standard; DNA; 711 BP.

XX

AC AAS53409;

XX

DT 13-FEB-2002 (first entry)

XX

DE Haemophilus influenzae DNA for cellular proliferation protein #191.

XX

KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
antibacterial; drug design.

XX

OS Haemophilus influenzae.

XX

PN WO200170955-A2.

XX

PD 27-SEP-2001.

XX

PF 21-MAR-2001; 2001WO-US009180.

XX

PR 21-MAR-2000; 2000US-0191078P.

XX

PR 23-MAY-2000; 2000US-0206848P.

XX

PR 26-MAY-2000; 2000US-0207727P.

XX

PR 23-OCT-2000; 2000US-0242578P.

XX

PR 27-NOV-2000; 2000US-0253625P.

XX

PR 22-DEC-2000; 2000US-0257931P.

XX

PR 16-FEB-2001; 2001US-0269308P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX

PI Yamamoto RT, Xu HH;

XX

XX WPI; 2001-611495/70.

XX

DR P-P8DB; AAU35550.

XX

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids.

XX

PS Claim 27; SEQ ID NO 7046; 51lpp; English.

XX

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes,
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic

CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences

XX
SQ Sequence 711 BP; 246 A; 113 C; 143 G; 209 T; 0 U; 0 Other;

Query Match 34.1%; Score 316.2; DB 4; Length 711;

Best Local Similarity 66.2%; Pred. No. 4.5e-94;

Matches 472; Conservative 0; Mismatches 238; Indels 3; Gaps 1;

QY 101 ATGCAGACCCCGCACATTTCTTATCGTTGAAGACGAGTTGGTAACACGCAACACGTTGAAA 160

Db 1 ATGACTACTCCAAAAATTTCTCGTTTGAAGATGAAATTTGTCACTCGAAATACGCTTAAA 60

QY 161 AGTATTTTCGAAGCGGAAGGCTATGATGTTTTCGAAGCGACAGATGGCGCGGAATGCAAT 220

Db 61 GGGATTTTGAAGCGGAAGGATATGATGTGTTGAAGCAGAAAAATGGTGTGTTGAAATGCAAT 120

QY 221 CAGATCCTCTCTGTAATATGACATCAACCTGCTGATCATGGATATCAATCTGCCGGGTAAG 280

Db 121 CATATTTGGCAATCATATTAATTTGTTGTTGATGGATATTAATTTACCAGGCAAA 180

QY 281 AACGGTCTTCTGTAGCGCGTGAATCTGCGAGCAGGCGAAATGTTGCGTTGATGTTCCCTG 340

Db 181 AACGGCTTATTTATTTGGCAAGAGAACTCCGTGAAGATTAAGCTTACCTCTTATTTTATA 240

QY 341 ACTGCCGTGACAAAGAGTCGATAAATTTCTGGCCTCGAATCCGTCGAGATCACTAC 400

Db 241 ACTGTCGAGATAAATGAAGTGGATAAATTTTAGGTTTGGAAATTTGGTCCGATGACTAT 300

QY 401 ATCAACCAACCGTTCAACCCGCGTGAATCTGTCGACGATTCGTGACGCAACCTACTGTCCT 460

Db 301 TTGACTAAGCCTTTTAACCTTAGAGAACTGACTATCCGCGCAGTAACTATTGTCATCGT 360

QY 461 ACCATGAATCTGGGTACTGTCTCAGCGAAGAACGTGCTAGCGTTGAAAGCTACAAAGTCAAT 520

Db 361 GCAATG---CCGCATCAAGAAAAAGAAATACATTTTGGTCGAGAAATCTATCGTTTAAAT 417

QY 521 GGTGGGAATCTGGACATCAACAGCGGTTCTGTTGATCGGCCCTGATGGCGAGCAGTACAG 580

Db 418 GGTGGAAATTTAGACCTAAATAGCCACAGTTTAAATTACACCAGAGAGCAAGAAATTCAAA 477

QY 581 CTGCGCGCAGCAGGATTCGCGCCATGCTTCACTTCTGTGAAACCCAGGCAAAATTCAG 640

Db 478 CTTCCTCGCAGTGAATTTCTGTCGAATGTTACATTTCTGTGAAATTCAGGAAATTCGAA 537

QY 641 TCCCGTGTCTGAACCTGCTGAAGAAATGACCGCCGCTGAGCTGAAACCGCACGACCGTACT 700

Db 538 ACGCGGAAGAAATGCTGAAAAAAATGACCGGACGAGAGTTAAAAACCTCAGGATCGTACC 597

QY 701 GTAGACGTGACGATCCGCCGTTATTCGTAAACATTTTCAATCTACGCCGATACGCGGNA 760

Db 598 GTAGATGTCACAATTCGACGATATCAGAAAAACATTTTGAAGATCATCCCAATCTCCAAT 657

QY 761 ATCATCGCCACCATTCACGGTGAAGGTTATCGCTTCTCGGTGATCTGGAAGA 813

Db 658 ATCATTATGACTATACATGGAGAGGCTATCGTTTTTGTGGAGATTAATGAATA 710

RESULT 13

ACA34287

ID ACA34287 standard; DNA; 711 BP.

XX

AC ACA34287;

XX

DT 19-JUN-2003 (first entry)

XX

DE Prokaryotic essential gene #15944.

XX

KW Antisense; ds; prokaryotic essential gene; cell proliferation;
drug design; gene.

XX

OS	Haemophilus influenzae.	
XX	WO200277183-A2.	
PN		
XX	03-OCT-2002.	
PD		
XX	21-MAR-2002; 2002WO-US009107.	
XX		
PF	21-MAR-2001; 2001US-00815242.	
XX	06-SEP-2001; 2001US-00948993.	
PR	25-OCT-2001; 2001US-0342923P.	
PR	08-FEB-2002; 2002US-00072851.	
PR	06-MAR-2002; 2002US-0362699P.	
XX		
XX	(ELIT-) ELITRA PHARM INC.	
XX		
XX	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KL, Zyskind JW;	
PI	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;	
PI	P-PSDB; ABU30417.	
XX		
DR	WPI; 2003-029926/02.	
DR	P-PSDB; ABU30417.	
XX		
XX	New antisense nucleic acids, useful for identifying proteins or screening	
PT	for homologous nucleic acids required for cellular proliferation to	
PT	isolate candidate molecules for rational drug discovery programs.	
XX		
PS	Claim 14; SEQ ID NO 22157; 1766pp; English.	
XX		
CC	The invention relates to an isolated nucleic acid comprising any one of	
CC	the 6213 antisense sequences given in the specification where expression	
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:	
CC	(1) a vector comprising a promoter operably linked to the nucleic acid	
CC	encoding a polypeptide whose expression is inhibited by the antisense	
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated	
CC	polypeptide or its fragment whose expression is inhibited by the	
CC	antisense nucleic acid; (4) an antibody capable of specifically binding	
CC	the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular	
CC	proliferation or the activity of a gene in an operon required for	
CC	proliferation; (7) identifying a compound that influences the activity of	
CC	the gene product or that has an activity against a biological pathway;	
CC	identifying a gene required for cellular proliferation or the biological	
CC	pathway in which a proliferation-required gene or its gene product lies	
CC	or a gene on which the test compound that inhibits proliferation of an	
CC	organism acts; (9) manufacturing an antibiotic; (10) profiling a	
CC	compound's activity; (11) a culture comprising strains in which the gene	
CC	product is overexpressed or underexpressed; (12) determining the extent	
CC	to which each of the strains is present in a culture or collection of	
CC	strains; or (13) identifying the target of a compound that inhibits the	
CC	proliferation of an organism. The antisense nucleic acids are useful for	
CC	identifying proteins or screening for homologous nucleic acids required	
CC	for cellular proliferation to isolate candidate molecules for rational	
CC	drug discovery programs, or for screening homologous nucleic acids	
CC	required for proliferation in cells other than <i>S. aureus</i> , <i>S. typhimurium</i> ,	
CC	<i>K. pneumoniae</i> or <i>P. aeruginosa</i> . The present sequence is one of the target	
CC	prokaryotic essential genes. Note: The sequence data for this patent did	
CC	not form part of the printed specification, but was obtained in	
CC	electronic format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences	
XX		
XX		
SQ	Sequence 711 BP; 246 A; 113 C; 143 G; 209 T; 0 U; 0 Other;	
	Query Match 34.1%; Score 316.2; DB 8; Length 711;	
	Best Local Similarity 66.2%; Pred. No. 4.5e-94;	
	Matches 472; Conservative 0; Mismatches 236; Indels 3; Gaps 1;	
QY	101 ATGCAGACCCCGCACATTCTTATCGTTGAAGACGAGTTGGTAACACGCAACACGTTGAAA	160
Db	1 ATGACTACTCCAAAATTCTCGTTTGAAGATGAAATTTGCTCCTCGAATACGCTTAAA	60
QY	161 AGTATTTTCGAAGCGGAGGCTATGATGTTTCGAAGCAGCAGATGGCGCGGAATGCAT	220
Db	61 GGGATTTTGAAGCGGAGGATATGATGTTTGAAGCAGAAAATGGTGTGAAATGCAT	120

QY	221	CAGATCTCTCTGAATATGACATCAACCTCGTGTGATCATGTATATCAATCTGCGGGTAAG	280
Db	121	CATATATTGGCAAAATCATATTAATTTGGTTGTGATGATATTAATTTACCAGCAAA	180
QY	281	AACGGCTCTTCTGTTAGCGCGTGAACCTGCGCAGCAGGCGAATGTTGCGTTGATGTTCTCG	340
Db	181	AACGGCTTTATTTATTTGGCAAGAGAACTCCGTGAAGAAATTAAGCTTACCTCTTTATTTT	240
QY	341	ACTGCGGTGACAAACGAGTGCATAAAATTTCTCGSCCTCGAAATCGGTGCGAGATGACTAC	400
Db	241	ACTGTCGAGATAATGAAGTGCATAAAATTTAGTTTGGAAATTTGGTCCGATGACATAT	300
QY	401	ATCACCAACCGTTTCAACCCCGCTGGAATCTGTCGACGATTCGTGACGCAACCTACTGTC	460
Db	301	TTGACTAAGCCCTTTTAACCCCTAGAGAACTGACTATCCGCGCACGTAATCTATTGTCAT	360
QY	461	ACCATGAATCTGGGTACTGTGTCAGCGAAGACGTCGTAGCGTTGAAAGCTACAAAGTTCAAT	520
Db	361	GCAATG---CCGCATCAAGAAAAAATAACATTTGGTCGAGAAATTCATCGTTTAAAT	417
QY	521	GGTTGGGAACCTGGACATCAACAGCCGTTTCGTTGATCGGCCCTGATGGCGAGCAGTACAA	580
Db	418	GGTTGGAAATTTAGACCTTAAATAGCCACAGTTTAATTTACACGAGAGGACAGAAATTC	477
QY	581	CTGCGCGCAGCGAGTTCCGCGCCATGCTTCACTTCTGTGAAAAACCCAGGCAAAATTCAG	640
Db	478	CTTCTCTCGCACTGAATTTGTCGTAATGTTTACATTTCTGTGAAATTCAGGAAATTC	537
QY	641	TCCCGTGTGAACTGCTGGAAGAAATGACCGCGCTGAGCTGAAACCGCAGCAGACCGTACT	700
Db	538	ACGCGCGAAGAAATTCGTAAGAAAAATGACCGCGCAGAGAGTTAAAAACCTTCAGGAT	597
QY	701	GTAGACGTGACGATCCGCGCTATTTCGTAAACATTTTGAATCTACGCGGATACGCCGAA	760
Db	598	GTAGATGTCACATTCGACGATTCAGAAACATTTTGAAGATCATCCCAATCTCCAAT	657
QY	761	ATCATCGCCACCATTCACGCTGAAGGTTTATCGCTTCTCGCGTGTGATCTGGAAGA	813
Db	658	ATCATATGACTATACATACGAGAGGCTATCGTTTGTGCGAGATATTGAATA	710

RESULT 14
AAS82422
ID AAS82422 standard; cDNA; 1110 BP.
XX
AC AAS82422;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #18226.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG18235.
XX

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2006, 13:37:39 ; Search time 4551 Seconds
(without alignments)
9530.131 Million cell updates/sec

Title: US-10-613-990A-31

Perfect score: 927

Sequence: 1 gtcattgtacgcgatcatg.....gtaactttcttactggttt 927

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_ges1:*
10: gb_ges2:*
11: gb_ges3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	481.2	51.9	772	10 CL681756	CL681756 PRI0132a
C 2	343	37.0	800	10 CL686716	CL686716 PRI0144d
C 3	114	12.3	114	2 BG834429	BG834429 352648 MA
C 4	112	12.1	841	2 BG367933	BG367933 HVSMBE1001
C 5	102	11.0	720	10 CNS00WTA	AL094012 Arabidops
C 6	99.6	10.7	768	10 CL685986	CL685986 PRI0142c
C 7	93.2	10.1	1193	9 BZ574214	BZ574214 msh2 3562
C 8	92.6	10.0	831	10 CL652351	CL652351 PRI0114c
C 9	92.6	10.0	1018	9 BZ574107	BZ574107 msh2 3514
C 10	89.8	9.7	1159	9 BZ559286	BZ559286 pacel2-164
C 11	89.6	9.7	1575	9 BZ574473	BZ574473 msh2 3684
C 12	88.6	9.6	1029	9 BZ551846	BZ551846 pacel1-60
C 13	84.8	9.1	732	10 CL944656	CL944656 O81FSB005
C 14	82.8	8.9	665	10 CW158500	CW158500 104_584.1
C 15	81.4	8.8	1189	9 BZ564167	BZ564167 pacel2-164
C 16	81.4	8.8	1356	9 BZ579121	BZ579121 msh2 6208
C 17	78.6	8.5	733	8 CV982963	CV982963 UMC-bmix
C 18	78.4	8.5	875	8 DR888097	DR888097 JGI_CUNH1
C 19	78.4	8.5	956	8 CV983114	CV983114 UMC-bmix
C 20	77.2	8.3	1347	9 BZ570026	BZ570026 msh2_1131
C 21	75.8	8.2	1166	9 BZ549106	BZ549106 pacel1-60
C 22	75.6	8.2	816	6 CB991419	CB991419 AGENCOURT

C 23	75.6	8.2	865	6 CB994454	CB994454 AGENCOURT
C 24	75	8.1	874	7 CN822966	CN822966 Oa_splbn
C 25	74.8	8.1	312	2 BE323316	BE323316 NF005H12P
C 26	74.4	8.0	544	9 BH017051	BH017051 TDGCS85TH
C 27	74.2	8.0	807	6 CB998053	CB998053 AGENCOURT
C 28	74.2	8.0	924	9 BZ549107	BZ549107 pacel1-60
C 29	73.6	7.9	911	8 DN640699	DN640699 UMC-bend
C 30	73.4	7.9	628	9 BZ551497	BZ551497 pacel1-60
C 31	72.6	7.8	667	2 BI263997	BI263997 NF092E12P
C 32	71	7.7	703	8 DN815843	DN815843 ACAC-aac1
C 33	71	7.7	768	7 CN821815	CN821815 Oa_splbn
C 34	71	7.7	871	7 CN824699	CN824699 Oa_splbn
C 35	70.2	7.6	847	9 BZ553164	BZ553164 pacel1-60
C 36	70	7.6	568	5 BU696722	BU696722 LL21n1207
C 37	69.4	7.5	555	7 CV122221	CV122221 Md1v4001j
C 38	69.4	7.5	594	7 CJ368177	CJ368177 CJ376658
C 39	69.4	7.5	697	7 CJ376658	CJ376658 pacel1-60
C 40	68.6	7.4	1263	9 BZ553503	BZ553503 pacel1-60
C 41	66	7.1	366	7 CO990188	CO990188 UMC-pd3ov
C 42	66	7.1	748	7 CN205135	CN205135 Tor5525 G
C 43	65.4	7.1	1274	9 BZ572156	BZ572156 msh2 2532
C 44	65.2	7.0	848	10 CL688368	CL688368 PRI0149b
C 45	64.8	7.0	867	10 CZ531116	CZ531116 SRAA-aac7

ALIGNMENTS

RESULT 1
CL681756/c

LOCUS

DEFINITION

PRI0132a All 2 - PRI0132a.BR (772) Mixed stage foamid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

CL681756 772 bp DNA linear GSS 09-JUL-2004

PRI0132a All 2 - PRI0132a.BR (772) Mixed stage foamid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.

CL681756 772 bp DNA linear GSS 09-JUL-2004

CL681756.1 GI:50188810

GSS.

Pristionchus pacificus

Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;

Neodiplogasteridae; Pristionchus.

1 (bases 1 to 772)

Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.

AppADB: an AcedB database for the nematode satellite organism

Pristionchus pacificus

Nucleic Acids Res. 32 (1), D421-D422 (2004)

14681447

Contact: Sommer RJ

Evolutionary Biology

Max-Planck-Institute for Developmental Biology

Spemannstr. 37-39, Tuebingen D-72076, Germany

Tel: 00497071601371

Fax: 00497071601498

Email: ralf.sommer@tuebingen.mpg.de

This library was generated at Caltech, Pasadena, USA and end

sequenced at Vancouver, Canada.

Seg primer: T7

Class: fosmid ends.

Location/Qualifiers

1. 772

/organism="Pristionchus pacificus"

/mol_type="genomic DNA"

/strain="California"

/db_xref="taxon:54126"

/clone_lib="Mixed stage fosmid library of P. pacificus var. California"

/note="Vector: pEpifos-5 Fosmid vector"

ORIGIN

Query Match 51.9%; Score 481.2; DB 10; Length 772;

Best Local Similarity 98.4%; Pred. No. 1.7e-140;

Matches 486; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 434 ATTCGTGACGCAACTACTCTCCGTACCATGAATCTGGGTACTGTCTAGCGAAGAACGT 493
DB 772 ATTCGTGACGCAACTCTGCTCCGTACCATGAATCTGGGTACTGTCTAGCGAAGAACGT 713
QY 494 COTAGCGTTGAAAGTACAAAGTTCAATGGTTGGGAATCTGGACATCAACAGCGCTTCGTTG 553
DB 712 COTAGCGTTGAAAGTACAAAGTTCAATGGTTGGGAATCTGGACATCAACAGCGCTTCGTTG 653
QY 554 ATCCGCCCTGATGGGACGAGTACAAAGCTGCGCGCAGCGAGTTCCGCGCATGCTTCAC 613
DB 652 ATCCGCCCTGATGGGACGAGTACAAAGCTGCGCGCAGCGAGTTCCGCGCATGCTTCAC 593
QY 614 TTCTGTGAAACCAGGCAAAATTCAGTCCCGTGTCTGAACTGCTGAAGAAATGACCGGC 673
DB 592 TTCTGTGAAACCAGGCAAAATTCAGTCTCGTGTCTGAACTGCTGAAGAAATGACCGGC 533
QY 674 COTGAGCTGAAACCCGACGACCGTACTGTAGACGTGACGATCCGCCGCTATTGCTAAACAT 733
DB 532 COTGAGCTGAAACCCGACGACCGTACTGTAGACGTGACGATCCGCCGCTATTGCTAAACAT 473
QY 734 TTGCAATCTACGCCGAGTACGCGGAATCATCGCCACCATTCACGGTGAAGTTATCGC 793
DB 472 TTGCAATCTACGCCGAGTACGCGGAATCATCGCCACCATTCACGGTGAAGTTATCGC 413
QY 794 TTCTCGGTGATCTCGAAGATTAAATCGGCTTTTACCACCGTCAAAAAAAGCGCGCTTTT 853
DB 412 TTCTCGGTGATCTCGAAGATTAAATCGGCTTTTACCACCGTCAAAAAAAGCGCGCTTTT 353
QY 854 AGCGCGGTTTTATTTTCAACCTTATTTCCAGATACGTAACCTCATCGTCCGTTGTAACT 913
DB 352 AGCGCGGTTTTATTTTCAACCTTATTTCCAGATACGTAACCTCATCGTCCGTTGTAACT 293
QY 914 TCTTTACTGGCTTT 927
DB 292 TCTTTACTGGCTTT 279

RESULT 2
LOCUS CL686716 800 bp DNA linear GSS 09-JUL-2004
DEFINITION PRI0144d_H03_2 - PRI0144d_BR (800) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic, genomic survey sequence.
ACCESSION CL686716
VERSION CL686716.1 GI:50195255
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
REFERENCE Srinivasan,J., Otto,G.W., Kahlow,U., Geiseler,R. and Sommer,R.J. AppADB: an AcedB database for the nematode satellite organism Pristionchus pacificus
JOURNAL Nucleic Acids Res. 32 (1), D421-D422 (2004)
PUBMED 1468147
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: raif.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end sequenced at Vancouver, Canada.
Seq primer: 17
Class: fosmid ends.
Location/Qualifiers
1..800
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"

FEATURES
Source

ORIGIN

Query Match 37.0%; Score 343; DB 10; Length 800;
Best Local Similarity 100.0%; Pred. No. 1e-96;
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCATGTTACGCCCATCATGTTAATTTGCAGCATGCATCAGGCAGGTTCAGGACTTTTGT 60
DB 458 GTCATGTTACGCCCATCATGTTAATTTGCAGCATGCATCAGGCAGGTTCAGGACTTTTGT 517
QY 61 ACTTCTCTGTTTCGATTTAGTTGGCAAATTTAGTAGCAAAATGATGACAGCCCGCACATTCT 120
DB 518 ACTTCTCTGTTTCGATTTAGTTGGCAAATTTAGTAGCAAAATGATGACAGCCCGCACATTCT 577
QY 121 TATCGTTGAAGACGAGTTGGTAAACACGCAACACGTTGAAAAGTATTTTCGAAGCGGAAGG 180
DB 578 TATCGTTGAAGACGAGTTGGTAAACACGCAACACGTTGAAAAGTATTTTCGAAGCGGAAGG 637
QY 181 CTATGATGTTTTCGAAGGACAGATGGCGGGAATGCATCAGATCCCTCTCTGAATATGA 240
DB 638 CTATGATGTTTTCGAAGGACAGATGGCGGGAATGCATCAGATCCCTCTCTGAATATGA 697
QY 241 CATCAACTGTTGATCATGGATATCAATCTCCGGGTAAGAACGCTCTCTGTTAGCGCG 300
DB 698 CATCAACTGTTGATCATGGATATCAATCTCCGGGTAAGAACGCTCTCTGTTAGCGCG 757
QY 301 TGAACCTCGCGAGCAGCGGAATGTTGGCTTGATGTTCTCTGACT 343
DB 758 TGAACCTCGCGAGCAGCGGAATGTTGGCTTGATGTTCTCTGACT 800

RESULT 3

LOCUS BG834429 114 bp mRNA linear EST 25-MAY-2001
DEFINITION BG834429 352648 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BG834429
VERSION BG834429.1 GI:14199698
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 114)

AUTHORS Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J., Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R., Quackenbush,J. and Keele,J.W.

TITLE Porcine gene discovery by normalized cDNA-library sequencing and EST cluster assembly

JOURNAL Mamm. Genome 13 (8), 475-478 (2002)

PUBMED 12226715

COMMENT

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4386

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 18

and -mismatch 12 options.

PCR Primers

FORWARD: AGGAAACAGTATGACCAT

BACKWARD: GTTTCACGTCACGACG

Plate: 113 row: K column: 19

Seq primer: ATTTAGGTGACACTATAG.

Location/Qualifiers

1..114

/organism="Sus scrofa"

/mol_type="mRNA"

/db_xref="taxon:9823"

/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1P1G"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

ORIGIN

Query Match 12.3%; Score 114; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 1.9e-24;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 559 CCTGATGCGGAGAGTCAAGCTCCGCGAGCGAGTTCGCGCCATGCTTCACTTCTG 618
Db 1 CCTGATGCGGAGAGTCAAGCTCCGCGAGCGAGTTCGCGCCATGCTTCACTTCTG 60
QY 619 TGAACCCAGCGCAAAATTCAGTCCCGTCTGAAGTCTGAAGAAATGACCGG 672
Db 61 TGAACCCAGCGCAAAATTCAGTCCCGTCTGAAGTCTGAAGAAATGACCGG 114

RESULT 4
BG367933 841 bp mRNA linear EST 22-OCT-2001
LOCUS HVSME10014N08f Hordeum vulgare 20 DAP spike EST library HVCDA0010
DEFINITION (20 DAP) Hordeum vulgare subsp. vulgare cDNA clone HVSME10014N08f,
mRNA sequence.

ACCESSION BG367933
VERSION BG367933.1 GI:13257032
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 841)
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D.,
Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W.,
Penton, R.D., Close, S.J., Oates, R. and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex 20 DAP spike cDNA library
Unpublished (2001)

COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 285
Seq primer: AATTAACTCTACTAAAGGG
High quality sequence stop: 388.

FEATURES

source
1..841
Location/Qualifiers
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Morex"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HVSME10014N08f"
/tissue_type="20 DAP spike"
/lab_host="SOLR"
/clone_lib="Hordeum vulgare 20 DAP spike EST library
HVCDA0010 (20 DAP)"
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
Plants were grown in the greenhouse at the University of
California, Riverside (Penton, SJ Close, TJ Close). Whole
spikes with awns trimmed were collected at 20 DAP
(Penton). Total RNA was prepared, poly(A) RNA was
purified, one primary unamplified cDNA library was made,
and 1 million pfu were in vivo excised to give pBluescript
SK(-) cDNA phagemids in the TJ Close lab at the University
of California, Riverside (Choi). Phagemids were plated and

picked at the Clemson University Genomics Institute (CUGI)
(Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
preparations, DNA sequencing and sequence analysis were
performed at CUGI (Wing, Yu, Frisch, Henry, Simmons,
Oates, Rambo, Main). The sequence has been trimmed to
remove vector sequence and contains a minimum of 100 bases
of phred value 20 or above. For more details on library
preparation and sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders/Alao
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

ORIGIN

Query Match 12.1%; Score 112; DB 2; Length 841;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 816 AATCGGCTTACCAACCGTCAAAAAACGGCGCTTTTAGCGCGCTTTTATTTTCAAC 875
Db 1 AATCGGCTTACCAACCGTCAAAAAACGGCGCTTTTAGCGCGCTTTTATTTTCAAC 60
QY 876 CTTATTTCCAGATACGTAACATCATCGCCGCTTGTAACCTCTTTTACTGCTTT 927
Db 61 CTTATTTCCAGATACGTAACATCATCGCTCGTGTGTAACCTCTTTTACTGCTTT 112

RESULT 5

CNS00WTA 720 bp DNA linear GSS 28-JUN-1999
LOCUS Arabidopsis thaliana genome survey sequence T7 end of BAC T12G15 of
DEFINITION TAMU library from strain Columbia of Arabidopsis thaliana, genomic
survey sequence.

ACCESSION AL094012
VERSION AL094012.1 GI:5295166
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 720)
Salanoubat, M., Choigne, N., Artiguenave, F., Brottier, P., Wincker, P.,
Samson, D., Saurin, W., Weissenbach, J. and Quetier, F.

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 720)
Genoscope.
AUTHORS Direct Submission

TITLE Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)

FEATURES

source
1..720
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="T12G15"
/clone_lib="TAMU"
/ecotype="Columbia"
/note="end : T7"

ORIGIN

Query Match 11.0%; Score 102; DB 10; Length 720;
Best Local Similarity 54.9%; Pred. No. 2.3e-20;
Matches 201; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
QY 108 CCCCGCATTCTTATCGTTCAAGACGAGTTGGTTAAACGCAACACGTTGAAAAGTATT 167
Db 350 CCATGAACGCTCTGCTCTGTCGAGACGACCGCGATCGTGATACCCCTGCTCAGGCT 409
QY 168 TCGAAGCGGAAGGCTATGATGTTTTCGAAGCGACAGATGGCGCGGAATGCATCAGATCC 227

```

Db      410 TGAAGACGAATGCTTCGAGGTGAAGACGCGCTCACGGGCGAGAGCGCGTGTGATGCT 469
QY      228 TCTCTGAATATGACATCAACCTGGTGATCATGGATATCAATCTGCCGGTGAAGAACGTC 287
Db      470 TTGGCGAAAAGCCCTTCGAGTTCGTGATCTCGACATCGGCTGCGCCGACATGACCGGCC 529
QY      288 TTCTGTTAGCGGTGAACCTGCGGAGCAGGCGAATGTTGCGTTGATGTTCTGACTGCC 347
Db      530 TTGACGCTTGCCGGAACCTCCGCGAAAGCTCGACCGTCCGCGTGTCTTCTTACCGTCC 589
QY      348 GTGACACGAAGTCGATATAAATTTCTCGGCTCGAAATCGGTGCAGATGACTACATCACCA 407
Db      590 GCGATGGGAGGTGACCGGATCTCTCGGCTCGAGCTGCGCGGGAGCACTACGTGACCA 649
QY      408 AACGTTTCAACCCCGGTGAATGACGATTCGTGACGCAACCTACTGTCCCGTACCATGA 467
Db      650 AGCCTTTTCAAGCCGCGTGAATCGTGGCAGCATTCGTGCCATCTCTGCGCGCTCGACCA 709
QY      468 ATCTCG 473
Db      710 GCCCG 715

```

```

RESULT 6
LOCUS   CL685986
DEFINITION CL685986 768 bp DNA linear GSS 09-JUL-2004
          PRI0142c_H02_2 - PRI0142c.BR (768) Mixed stage fosmid library of P.
          pacificus var. California Pristionchus pacificus genomic, genomic
          survey sequence.
ACCESSION CL685986
VERSION   CL685986.1 GI:50194404
KEYWORDS  GSS.
SOURCE    Pristionchus pacificus
          Pristionchus pacificus
          Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
          Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 768)
AUTHORS   Srivivasan,J., Otto,G.W., Kahlow,U., Geiseler,R. and Sommer,R.J.
TITLE      AppADB: an AcedB database for the nematode satellite organism
          Pristionchus pacificus
JOURNAL    Nucleic Acids Res. 32 (1), D421-D422 (2004)
PUBMED     14681447
COMMENT    Contact: Sommer RJ
          Evolutionary Biology
          Max-Planck-Institute for Developmental Biology
          Spemannstr. 37-39, Tuebingen D-72076, Germany
          Tel: 00497071601371
          Fax: 00497071601498
          Email: ralf.sommer@tuebingen.mpg.de
          This library was generated at Caltech, Pasadena, USA and end
          sequenced at Vancouver, Canada.
          Seq primer: T7
          Class: fosmid ends.

```

```

FEATURES             Location/Qualifiers
     source            1..768
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Mixed stage fosmid library of P. pacificus
                     var. California"
                     /note="Vector: pEpifos-5 Fosmid vector"

```

```

ORIGIN
Query Match      10.7%; Score 99.6; DB 10; Length 768;
Best Local Similarity 54.7%; Pred. No. 1.3e-19;
Matches 198; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY      102 TCGAGACCCCGCACATCTTATTCGTTGAAGACGAGTTGGTATACACGCAACGCGTTGAAAA 161
Db      243 TGATGCCACATCACATTTGTTATTTGTTGAAGATGAGCCGGTTACCAGCGCGCATTAACAAT 302

```

```

QY      162 GTATTTTCGAAGCGGAGGCTATGATGTTTTCGAAGCGACAGATGGCGCGGAAATGCATC 221
Db      303 CCTACTTACTCAGAGGGGTATACCGGTTTCGGTTACGGCGAGCGGTGCCGGGTACCGG 362
QY      222 AGATCCTCTCTGAATATGACATCAACCTGGTGATCATGGATATCAATCTGCCGGTAAAGA 281
Db      363 AAATATATGCAGAATCAGTCGGTAGATTTAATTCCTGCTGGATATCAACTTACCAGATGAAA 422
QY      282 ACGTCTTCTGTTAGCGGTGAACCTGCGGAGCAGGCGAATGTTGCGTTGATGTTCTCTGA 341
Db      423 ATGCGCTGATGTTTAAACCGCGCCCTCGGAGAACCTCAACCGTGGGGATTTATCTGGTTA 482
QY      342 CTGCGCGTGACACGAAGTCGATAAAATTTCTCGGCTCGAAATCGGTGCAGATGACTACA 401
Db      483 CCGAGCGAGGATCGGATTCACCGTATTTGTTGGCTGGAAATGGCGGCAACGATTAAG 542
QY      402 TCACCAACCGTTCAACCCCGGTGAACCTGACGATTCGTGACGCAACCTACTGTCCCGTA 461
Db      543 TCACCAACCGCTGGAACCTGCGGCAACTGGTAGTACGGGTGAAAAATCTGCTCTGGCGAA 602
QY      462 CC 463
Db      603 TC 604

```

```

RESULT 7
LOCUS   BZ574214/c
DEFINITION BZ574214 1193 bp DNA linear GSS 17-DEC-2002
          msh2_3562.y3 msh Pseudomonas aeruginosa genomic clone msh2_3562,
          genomic survey sequence.
ACCESSION BZ574214
VERSION   BZ574214.1 GI:27209275
KEYWORDS  GSS.
SOURCE    Pseudomonas aeruginosa
          Pseudomonas aeruginosa
          Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
          Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1193)
AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
          Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE      Whole-Genome-Sequence variation among multiple isolates of
          Pseudomonas aeruginosa library
JOURNAL    J. Bacteriol. (2002) In press
PUBMED
COMMENT    Contact: Chris K. Raymond
          Genome Center
          University of Washington
          Box 352145, Seattle, WA 98105-2145, USA
          Tel: 2062216954
          Fax: 2066857244
          Email: craymond@u.washington.edu
          Class: shotgun.

```

```

FEATURES             Location/Qualifiers
     source            1..1193
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="M5H"
                     /db_xref="taxon:287"
                     /clone_lib="msh2_3562"
                     /clone_lib="msh"
                     /note="Environmental isolate. Whole genomic shotgun
                     library."

```

```

ORIGIN
Query Match      10.1%; Score 93.2; DB 9; Length 1193;
Best Local Similarity 53.8%; Pred. No. 1.6e-17;
Matches 191; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY      112 GCACATTCCTTATCGTTGAAGACGAGTTCGTTAACACGCAACACGTTGAAAGATTTTTCGA 171
Db      579 GCGCATCTGTTGATCGAAGACGATACGAACGCGCGAGTACTCTGAAGAGGGGCTCGG 520
QY      172 AGCGGAAGGCTATGATGTTTTCGAAGCGACAGATGGCGCGGAAATGCATTCAGATCCTCTC 231

```


Db 519 CGAGTCGGGCTATGCGGTGCACTGGTGGCAGCACGGCGCCGACGGTCTCTACCTCGCGCT 460

Qy 232 TGAATATGACATCAACCTCGGTGATCATGTGATATCAATCTGCGGGTAAAGACGGTCTTCT 291

Db 459 GGAGAACCGCTTCGACCTGGTGGTCTCGACTGATGCTGCGCGCTGGACGGTTGGCA 400

Qy 292 GTTAGCGGTGAACCTCGCGGACGAGCGAATGTTGCGTTGATGTTCTTCACTGACCTGGCGTGA 351

Db 399 GATCATGGAAGTGTGCGCAAGACGATGTCGGGTGCTCTTCTCCTCACCGCCGCGA 340

Qy 352 CAACGAAGTCGATAAAATTTCTCGGCTCGAAATCGGTGCGAGATCACTACATCACCACAC 411

Db 339 CCGCTGCGAAGCCGATATCCGCGGCTCGACTGCGGTGCTGACGACTACCTGGTGAACC 280

Qy 412 GTTCAACCGCGTGAACCTGACGATTGCTGCGACGCAACCTACTGTCCCGTACCATG 466

Db 279 CTCTCTTCCACCGAGTGTCTGCTGCTATCGTACCTCTGCTGCGCGCGGGTG 225

RESULT 8
CL652351/c
LOCUS
DEFINITION
CL652351 831 bp DNA linear GSS 09-JUL-2004
PRI0114c_E11 - PRI0114c_B21 (831) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE
AUTHORS
TITLE
Srinivasan, J., Otto, G.W., Kahlow, U., Geisler, R., and Sommer, R.J.
AppADB: an AcedB database for the nematode satellite organism
Pristionchus pacificus
Nucleic Acids Res. 32 (1), D421-D422 (2004)
14681447
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@uebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.

FEATURES
source
1. .831
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/note="Vector: pEpifos-5 Fosmid vector"

ORIGIN
Query Match 10.0%; Score 92.6; DB 10; Length 831;
Best Local Similarity 54.6%; Pred. No. 2.3e-17;
Matches 185; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

Qy 102 TCAGACCCCGACATCTTATCTGTTGAAGACGAGTTGTTACACGCAACAGTTGAAAA 161

Db 340 TGATGCCACATCACATTTGTTATGTTGAAGATGAGCCGGTTACCCAGCGCGATTTACAA 281

Qy 162 GTATTTTCGACGGAAGCTATGATGTTTTCGAAGCGACAGATGCGCGGAAATGCAATC 221

Db 280 CCTACTTCACTCAGAGGGGTATACCGTTTCGTTACGCGAGCGGTGCCGGGTACGGG 221

Qy 222 AGATCCTCTCTGAATATGACATCAACCTGGTGTATCATGATATCAATCTGCGGGTAAGA 281

Db 220 AAATATTCAGAAATCAGTCGGTAGATTAAATTCCTGCTGATATCAATTCACCGATGAAA 161

Qy 282 ACGTCTTCTCTAGCGGTGAATCGCGGACGAGCGGAATGTTGCGTTGATGTTCTCTGA 341

Db 160 ATGGCTCATGTTTAAACCGCGCCCTGCGAGAACGCTCAACGCTGGGGATTTCTGGTTA 101

Qy 342 CTGGCCGTGACACGAAGTCGATAAAATTTCTCGGCTCGAAATCGGTGCGAGATGACTACA 401

Db 100 CCGGACGCGAGATCGGATTGACCGTATTGTTGGAAGGGAATGCGCGCAGACGATTACG 41

Qy 402 TCACCAACCGTTCAACCCGCGTGAACCTGACGATTCTGTTG 440

Db 40 TCACCAACCGCTAAAGTGGCGAACTGGTAGTACGGG 2

RESULT 9
BZ574107/c
LOCUS
DEFINITION
BZ574107 1018 bp DNA linear GSS 17-DEC-2002
msh2_3514.x1 msh Pseudomonas aeruginosa genomic clone msh2_3514,
genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Pseudomonas aeruginosa
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES
source
1. .1018
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="M5H"
/db_xref="taxon:287"
/clone_lib="msh2_3514"
/clone_lib="msh"
/note="Environmental isolate. Whole genomic shotgun
library."

ORIGIN
Query Match 10.0%; Score 92.6; DB 9; Length 1018;
Best Local Similarity 53.8%; Pred. No. 2.4e-17;
Matches 191; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

Qy 112 GCACATCTTATCGTTGAAGACGAGTTGTTAAACGCAACACGTTGAAAAATTTTCGA 171

Db 560 GCGCATCTCTGTTGATCGAAGACGATACGAAGACCGCGAGTACCTCGAAGAGGGCGCTCGG 501

Qy 172 AGCGGAAGGCTATCATGTTTTCGAAGCGACAGATGCGCGGAATGTCATCATGCTCTC 231

Db 500 CGATTCGGGCTATCGGGTCTGACTGCTGTCGACACGCGCGCCGACGGTCTCTACCTGGCGCT 441

Qy 232 TGAATATGACATCAACCTCGGTGATCATGGATATCAATCTGCGGGTAAAGACGGTCTTCT 291

Db 440 GGAGAACCGCTACGACCTCTGGTGGTCTCTCGAGTATGCTGCCCGCCCTGGACGGTTGCA 381

Qy 292 GTTAGCGGTGAACCTGCGCGAGCGAGCGAATGTTGCGTTGATGTTCTCTGACTGGCGGTGA 351

Db 380 GATCATGAAGTGTGCGCAAGAACGACGATGTGCGGTGCTCTTCTCTCACGCCCGCA 321

```

QY 352 CAACGAGTCGTATAAATCTCGGCTCGAATCGGTGAGTACTACATCACCACACC 411
Db 320 CAGGCTGCAAGACCGTATCCGGCGCTCGAACTGGGTGCTGACGACTACCTGGTGAACACC 261
QY 412 GTTCAACCGCGGTGAACCTGACGATTCTGTCGACGCAACCTACTGTCTCCGTACCATTG 466
Db 260 CTTCTCCTTCCACGAGTTGCTGCTGGTATCGTATACCCTGCTGCGCGCGGGGCTG 206

RESULT 10
LOCUS BZ559286 1159 bp DNA linear GSS 17-DEC-2002
DEFINITION pacs2-164_1372.b1 pacs2-164 Pseudomonas aeruginosa genomic clone
ACCESSION BZ559286
VERSION BZ559286.1 GI:27175379
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1159)
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
    source
        1..1159
            Location/Qualifiers
                /organism="Pseudomonas aeruginosa"
                /mol_type="genomic DNA"
                /strain="2-164"
                /db_xref="taxon:287"
                /clone="pacs2-164_1372"
                /clone_lib="pacs2-164"
                /note="clinical isolate 2-164 Whole genomic shotgun
                    library."
ORIGIN
    Query Match 9.7%; Score 89.8; DB 9; Length 1159;
    Best Local Similarity 55.2%; Pred. No. 1.9e-16;
    Matches 197; Conservative 0; Mismatches 157; Indels 3; Gaps 1;
QY 243 TCAACCTGCTGATCATGGATATCAATCTGCGGGTAAGAACGGTCTTCTGTTAGCGGTG 302
Db 312 TCGAGCTGCTGCTGCTGGACATCCGCTGCGCGCAAGGACGGCTGACCTGACCCGCG 371
QY 303 ACTGCGGAGCAGGCGAATGTGCTTATGTTCTGACTGCGCGTGNACAGAACTCG 362
Db 372 AGCTGCGGTGGGCTCCGAGGTGGGGATCATCTGATCACCGGGGCAACGACGAGATCG 431
QY 363 ATATAAATCTCGGCTCGAAATCGGTGACAGATGACTATACATACCAAAACGGTTCAACCCGC 422
Db 432 ATCGCATCGTCGGCTGGAGTGGGCGCGACGATTAAGCGCTGAACCCGCG 491
QY 423 GTGAATGACGATTCGTGCACGCAACCTACTCTGCCGTACCAATGAATCTGGGTACTGTGA 482
Db 492 GCGAACTGCTGTCGCGGGGGAAGAAC---CTGATCCGCGGTGGCCATGCCCCAGGCCA 548
QY 483 GCGAAGAACGTCGTAGCTGTTGAAGCTACAAGTTCAATGGTTGGAACTGGACATCAACA 542
Db 549 GCGCGGGCCCCCGCGGAGGCCCTCAGGCAGTTTGGCGACTGGCTGTGGACCGCGACC 608
QY 543 GCCGTTTGTGATCGGCCCTGTATGGCGAGCAGTACAAGCTGCGCGGACGAGGATTCC 599

```

```

Db 609 GGAGCCGCTGATGACACCGGCAACGAACCCCTGCTCAACCATGGCGAATTCC 665

RESULT 11
LOCUS BZ574473 1575 bp DNA linear GSS 17-DEC-2002
DEFINITION meh2_3684.x1 meh Pseudomonas aeruginosa genomic clone meh2_3684,
genomic survey sequence.
ACCESSION BZ574473
VERSION BZ574473.1 GI:27209534
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1575)
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
    source
        1..1575
            Location/Qualifiers
                /organism="Pseudomonas aeruginosa"
                /mol_type="genomic DNA"
                /strain="MSH"
                /db_xref="taxon:287"
                /clone="meh2_3684"
                /clone_lib="meh"
                /note="Environmental isolate. Whole genomic shotgun
                    library."
ORIGIN
    Query Match 9.7%; Score 89.6; DB 9; Length 1575;
    Best Local Similarity 51.3%; Pred. No. 2.5e-16;
    Matches 234; Conservative 0; Mismatches 219; Indels 3; Gaps 1;
QY 345 GCCGTGACACGAAAGTCGATATAAATTTCTCGGCTCGAAATCGGTGCAGATGACTACATCA 404
Db 88 GCCCCACACGACGAGATCGATCGCATCGTCGGCTGGAGTGGCGGCCGACGATTACGTGA 147
QY 405 CCAAACCGTTCAACCCCGGTGAACCTGACGATTTCGTGCACGCAACCTACTGTCCCGTACCA 464
Db 148 TCAAGCCCTGAAACCCGCGCAACTGGTGTGCGGGGCGAAGAG---CTGATCCGCGGG 204
QY 465 TGAATCTCGGTACTGTTCAGCGAAGAACCGTCGTAGCGTTTGAAGGTATCAAGTTCAATGGTT 524
Db 205 TGGCCATGCCGACGCGCAGCCGCGCCCGCCGCGAGCCCTCAGGCAGTTTCGGGACT 264
QY 525 GGGAACTCGACATCAACAGCGCTTGTGTTGATCGGCGCTGTATGCGGAGCAGTACAAGCTGC 584
Db 265 GGCTGTCGACCGCGACCGGCGCGCTGTATCGACACGCGGGCAACGAAACCTGCTCA 324
QY 585 CGCGCAGCGAGTTCGCGGCCATGCTTCACTTCTGTGAAACCCAGCAAAATTCAGTCCC 644
Db 325 CCCATGGCGAATTCCAACTGCTGGCGCTTTCCTCGGCAACAGCAGGCATACCTTGAGCC 384
QY 645 GTGCTGAATCTGTGAAGAAATGACCGGCGGTGAGCTGAAACCGCACGACCGTACTGTAG 704
Db 385 GCGACCACTTGTATGGACAGATCCGCAACCGCGAGTGGCTGCCAGCGATCGCTCCATCG 444
QY 705 ACGTGAAGATCGCGGTATTGTGAACATTTGGAATCTACGCGGATACGCGGGAATCA 764
Db 445 ACGTGTGTGCGCGCGCTGCGTGCAGAGCTGCGCAGACGCCCGCGCAACCGCAACTGA 504

```



```
QY 333 TGTTCCTGACTGGCGGTGACACGAGTCGATAAAATTTCTGGCCTCGAAATCGGTGCAG 392
Db 251 TATCTCTACCCGCAAGGCGGAGGACGTGGATCGCATCGTGGGCTGGAAAGTGGGTGCCG 310
QY 393 ATGACTATACATCAACAAACCGTTCAACCCGCGTGAACCTACGATTCGTGACGCAACCTTAC 452
Db 311 ACGACTACTGGGCAAGCCCTTCATCCCGGAGCTGCTGGCGGCATCCACGCCGTGC 370
QY 453 TGTCCCGTACCATGAATCTGGGTACTGTGACGCAAGAACGTGCTAGCTTGAAGCTACA 512
Db 371 TGGCGCGCGCCGCGCAGGAGGACCGGGCGCCCTCGGGCGACAAACGAGGTGCTGA 430
QY 513 AGTTCAATGGTGGGNACTGGACATCAACAGCCGTTCTGTGATCGGCCCTGATGGCGAGC 572
Db 431 CCTTCGGCCCCCTTACCTTCGACCTGGGCAACGCGCCCTTGCAAGA---ACGGCGAGG 487
QY 573 AGTACAAGCTGCGCGCGACGAGTTCCGCGCCATCTTCACTTCTGTGAAACCCAGGCA 632
Db 488 AGCTGCCCTTACACACCGCGGAATTCGCCATGCTCAAGGCCCTGTGGCGCCACCGCGCC 547
QY 633 AAATTCAGTCCGCTGTAACCTGTGTAAGAAATGACCGCGCGTGAAGCTGAAACCGCAAG 692
Db 548 AGCCGCTGTGGCGCAAGAAAGCTGGCCCTGCTGGCCCGCGCGCGAGTTTCGAGCCCTTCG 607
QY 693 ACCGTACTGTAGACGTGACGATCCCGCGTATTCGTAAACATTTTCAATCTACCGCGATA 752
Db 608 ACCGAGCGCTGACGTGACGATCTCGCGCTCGCGAAGCTGTGAGTGAGATGCGCGCG 667
QY 753 CCGCGGAAATCATCCGCCACCATTCACGGTGAAGTTTATCGTTTC 796
Db 668 CTCCCCGCTACATCCAGACCGTCTGGGGCGTGGGTACGTGTTTC 711
```

```
RESULT 14
CW158500/c 665 bp DNA linear GSS 29-OCT-2004
LOCUS 104_564_11148854_148_36401_059 Sorghum methylation filtered library
DEFINITION (LibID: 104) Sorghum bicolor genomic clone 11148854, genomic survey
sequence.
ACCESSION CW158500
VERSION CW158500.1 GI:54851047
KEYWORDS GSS.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 665)
AUTHORS Bedell J.A., Rudiman, M.A., Nunberg, A., Citek, R.W., Robbins, D.,
Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K.,
McMenamy, J., Smith, M., Holeman, H., Roe, B.A., Wiley, G., Korff, I.F.,
Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddalo, J.A. and
Martienssen, R.A.
TITLE Sorghum genome sequencing by methylation filtration
JOURNAL PLoS Biol. 3 (1), e13 (2005)
PUBMED 15660154
COMMENT Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: 564 row: f column: 14
Seq primer: SWfor Forward
Class: methylation filtered
High quality sequence stop: 665.
FEATURES
Location/Qualifiers
1..665
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATx623"
/db_xref="taxon:4558"
```

```
/clone="11148854"
/clone_lib="Sorghum methylation filtered library (LibID:
104)"
/note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA
prepared from purified nuclei was randomly sheared,
end-repaired, size fractionated to enrich for the 0.5 to 5
kb fraction, ligated into HincII-digested pBCSK(-) vector
and electroporated into E. coli cells. This is a
methylation filtered library."
```

```
ORIGIN
Query Match 8.9%; Score 82.8; DB 10; Length 665;
Best Local Similarity 52.3%; Pred. No. 2.7e-14;
Matches 183; Conservative 0; Mismatches 167; Indels 0; Gaps 0;
QY 110 CCGCACATCTTATTCGTTGAAGACGAGTTGTGAACACGACACGTTGAAAGTATTTTC 169
Db 471 CCCACATCTCTGGTCTGATGACACACCGCCCTGCGCGAACTGCTCGCAAGTATCTC 412
QY 170 GAAGCGGAGCTATGATGTTTTCGAGCGACAGATGGCGGAAATGCATCAGATCCTC 229
Db 411 GCCGCGAACGCTTCTGCTGGCCACGCAAGGACGCGCGGAGCGCGGCCAAGCTG 352
QY 230 TCTGAATATGATCAACCTGCTGATCATGATATCAATCTGCCGGGTAAAGACGCTTT 289
Db 351 GCCGGCTCGCTTCGACCTGCTGCTGCATCATGATGCCGGCGAGACCGGGCTG 292
QY 290 CTGTTAGCGCGTAACTGCGGAGAGCGAATGTTGGTGTGATGTTCTGACTGGCGGT 349
Db 291 GAGCTGACCGAATCGCTGCGCGGAGCGACCTGCCCATCTGCTGCTGACCGCGG 232
QY 350 GACACGAGTCGATATAAATTTCTGGCTCCGCAATCGTGCAGATGACTACATCACCATA 409
Db 231 GCGGAGCGGACGACCGCATCGCCGGTTGGAGCGCGGGCGGACGATTCCTGCCAAG 172
QY 410 CCGTTCAACCGCGTGAACCTGATTCGTCGACGCAACCTACTGTCTCCG 459
Db 171 CCCTTCAACCGCGGAGCTCTGCTGCGCATCACTCGATCCTGCGCG 122
```

```
RESULT 15
BZ564167/c 1189 bp DNA linear GSS 17-DEC-2002
LOCUS pac82-164_4617.xl pac82-164 pseudomonas aeruginosa genomic clone
DEFINITION pac82-164_4617, genomic survey sequence.
ACCESSION BZ564167
VERSION BZ564167.1 GI:27189394
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1189)
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
Location/Qualifiers
1..1189
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pac82-164_4617"
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2006, 13:37:39 ; Search time 215 Seconds
(without alignments)
7664.184 Million cell updates/sec

Title: US-10-613-990A-31
Perfect score: 927
Sequence: 1 gtcattgtacgcgatcatg.....gtaactcttactggcttt 927

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/FCUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	577.2	62.3	726	3	US-09-489-039A-1812
2	450.8	48.6	726	3	US-09-543-681A-2033
3	319.4	34.5	1830121	3	US-09-557-884-1
4	319.4	34.5	1830121	3	US-09-643-990A-1
5	319.4	34.5	1830121	3	US-10-158-865-1
6	126	13.6	837	3	US-09-252-991A-16417
7	126	13.6	2430	3	US-09-252-991A-16171
8	126	13.6	3675	3	US-09-252-991A-16309
9	120.6	13.0	990	3	US-09-252-991A-11517
10	120.6	13.0	2166	3	US-09-252-991A-11722
11	113.8	12.3	2004	3	US-09-252-991A-11865
12	107.4	11.6	1018	3	US-09-418-980-1
13	98.6	10.6	729	3	US-09-634-238-133
14	97.2	10.5	762	3	US-09-252-991A-6307
15	96.2	10.4	705	3	US-09-134-001C-226
16	94.6	10.2	702	3	US-09-710-279-479
17	94.6	10.2	3389	3	US-09-710-279-3914
18	92.6	10.0	792	3	US-09-252-991A-10543
19	92.6	10.0	804	3	US-09-252-991A-10323
20	91.4	9.9	699	3	US-09-925-637-45
21	91.4	9.9	702	3	US-09-082-077-2
22	91.4	9.9	3731	3	US-09-082-077-1
23	91.4	9.9	9425	3	US-08-956-171E-87
24	91.4	9.9	9425	3	US-08-781-986A-87

ALIGNMENTS

RESULT 1
US-09-489-039A-1812
; Sequence 1812, Application US/09489039A
; Patent No. 6610936
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1812
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1812

Query Match 62.3%; Score 577.2; DB 3; Length 726;
Best Local Similarity 87.2%; Pred. No. 3.3e-190;
Matches 633; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

Qy	92	GTAGCAAAATGTCAGACCCCGCACATTTCTTATCGTTGAAGACGAGTTGGTAACACGCAAC	151
Db	1	GTAGCAAAATGTCAGACCCCGCACATTTCTTATCGTTGAAGACGAGTTGGTAACACGCAAC	60
Qy	152	ACGTTGAAAAGTATTTTCGAAGCGGAAGCTATGATGTTTTCGAAGCGACAGATGCGCG	211
Db	61	ACGTTGAAAAGTATTTTCGAAGCGGAAGCTATGATGTTTTCGAAGCGACAGATGCGCG	120
Qy	212	GAATGTCATCAGATCCTCTCTGTAATATGACATCAACCTGGTGATCATGATATCATCTG	271
Db	121	GAATGTCATCAGATCCTCTCTGTAATATGACATCAACCTGGTGATCATGATATCAACCTG	180
Qy	272	CGGGTGAAGAAGCGTCTTCTGTAGCGCGTGAACCTGCGGACGACGCGAATGTTGGGTTG	331
Db	181	CGGGTGAAGAAGCGTCTTCTGTAGCGCGTGAACCTGCGGACGACGCGAATGTTGGGTTG	240
Qy	332	ATGTTCTTGACTGCGCGTGACAAACGAGTCGATAAAATTTCTCGGCTCGAAATCGGTGCA	391
Db	241	ATGTTCTTAAACGCGCGCGGACAAACGAGTCGATAAGATCTTGGCTCGAAATCGGCGCT	300
Qy	392	GATGACTATCATCACCACAAACCGTTTCAACCGCGGTGAACGATTCGTGTCAGCAACCTA	451
Db	301	GACGACTATATCATTAAACCGTTTAAACCGCGGTGAACCTGACTATCCGCGCGCGCAACCTG	360

QY 452 CTGTCCTCCGTACCATGAATCTGGTACTGTCTGACGGAAGAACGTCGTAGCGTTGAAAGCTTAC 511
DB 361 CTCTCCCGTACCATGAACCTCGGTACCGTGAGCGGAAGAGCGTCGCAGCGTGGAAAGCTTAC 420
QY 512 AAGTTCAATGGTTGGGAACTGACATCAACAGCGGTTCTGTTGATCGGCCCTGTAGTGGCGAG 571
DB 421 AAGTTCAACGGCTGGGAACCTGATATCAACAGCCGTTCTCCTGGTTAGCCCGGAAACCGGCGAA 480
QY 572 CAGTACAAAGCTGCCCGCAGCGAGTTCCGCGCATCTTCACTTCTGTGAAACCCAGGC 631
DB 481 CAGTACAAAGCTGCCCGCAGCGAGTTCCGCGCATCTTCACTTCTGTGAAACCCAGGC 540
QY 632 AAAATTTCAGTCCCGTCTGAACTGCTGAAGAAATGACCGCGCTGAGCTGAACCGCAC 691
DB 541 AAAATTTCAGTCCCGTCTGAACTGCTGAAGAAATGACCGCGCTGAGCTGAACCGCGCAT 600
QY 692 GACCGTACTGTAGACGTGACGATCCGCGTATTTCGTAACATTTTCGAATCTACGCCGGAT 751
DB 601 GACCGTACTGTAGACGTGACGATCCGCGTATTTCGTAACATTTTCGAATCTACGCCGGAT 660
QY 752 AGCCGGAATATCATCCCAACCATTCACCGTGAAGGTTATCGGTTCTCGCGGTGATCTGGAA 811
DB 661 ACCCGGAATATCATCCCAACCATTCACCGTGAAGGTTATCGGTTCTCGCGGTGATCTGGCA 720
QY 812 GATTAA 817
DB 721 GAATAA 726

RESULT 2

US-09-543-681A-2033

; Sequence 2033, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 2033

; TYPE: DNA

; LENGTH: 726

; ORGANISM: Proteus mirabilis

US-09-543-681A-2033

Query Match 48.6%; Score 450.8; DB 3; Length 726;
Best Local Similarity 76.3%; Pred. No. 3.8e-146;
Matches 554; Conservative 0; Mismatches 172; Indels 0; Gaps 0;
QY 92 GTAGCAAAATATGCAACCGCCGACATCTTATCGTTGAAGACGAGTGTGAACACGCAAC 151
DB 1 GTAGCAAAATATGCAACCGCCGACATCTTATCGTTGAAGACGAGTGTGAACACGAGTGTACTCGTAAT 60
QY 152 ACGTTGAAAGTATTTTTCGAAGCGGAAGCTATGATGTTTTCGAAGCGACAGATGCGCG 211
DB 61 ACCCTGAAAGCATATTCGAAGCTGAAGGTATATCGTACACGAAGCCACTGATGGCAAC 120
QY 212 GAATGCAATCAGATCGTCTCTGAATATGACATCAACCTGGTGATCATGTGATATCAATCTG 271
DB 121 GAGATGCAATATTTCTGTCCGACCATGATATCAATCTGGTCAATTTATGATATTAATCTT 180
QY 272 CCGGTTGAAGACGCTCTTCTGTTAGCGGTGAACCTGCGGACGACGCAATCTTCCGTTG 331
DB 181 CTGGTAAATATGGTCTTCTTATAGCCGTGAATTAACGTGAACAGGTAGTGTTCATTA 240
QY 332 ATGTTCTGACCTGGCGGTGACACGAAGTCGATAAAATTTCTCGGCTCGAAATCGGTGCA 391
DB 241 ATGTTCTTAACAGGTGCTGATATGAAGTTGATAAAATCTTAGGCTTTGAAATTCGTGCC 300

QY 392 GATGACTACATCACCBAACCGTTTCAACCCGCGTGAACTGACGATTCGTGACGCAACCTA 451
DB 301 GATGATTACATCACTAAACCATTTAATCCTCGTGAAATTAACATATCGGTGCTCGTAACCTTA 360
QY 452 CTGTCCTCCGTACCATGAATCTGGTACTGTCTGACGGAAGAACGTCGTAGCGTTGAAAGCTTAC 511
DB 361 TTGTCAAGCATATGAATTTAGCGAATGGCACAGAGCGTCGTTTGTGAAAGCTTAT 420
QY 512 AAGTTCAATGGTTGGGAACTGGAACATCAACAGCCGTTCTGTTGATCGGCCCTGTAGTGGCGAG 571
DB 421 AAATTTAATGGTTGGGAGCTAGATATTAATAGTCGCTCTCTATTATTAGCCCTACAGGTGAA 480
QY 572 CAGTACAAAGCTGCCCGCAGCGAGTTCCGCGCATCTTCACTTCTGTGAAACCCAGGC 631
DB 481 CAGTATAAATTTACCTCGTAGTGAGTTTCGTCGCGATTTTACATTTCTGCGAAACCCAGGA 540
QY 632 AAAATTTCAGTCCCGTCTGAACTGCTGAAGAAATGACCGCGCTGAGCTGAAACCCGAC 691
DB 541 AAAATCCAACTCGTCGAGATTTACTGAAGAAATGACGGGTGCTGAATTTAAACCTCAT 600
QY 692 GACCGTACTGTAGACGTGACGATCCGCGTATTTCGTAACATTTTCGAATCTACGCCGGAT 751
DB 601 GATCGTACTGTAGACGTGATCCCATTCGTCGTATTCGTAACACATTTTGAATCAACCCCTGAT 660
QY 752 ACGCGGAATATCATCCCAACCATTCACGGTGAAGGTTATCGGTTCTCGCGGTGATCTGGAA 811
DB 661 ACCTTGAGATTTATCGCCACTATCCATCGTGAAGGTTATCGTTTCTGTTGGTGAATTTAGAC 720
QY 812 GATTAA 817
DB 721 GATGA 726

RESULT 3

US-09-557-884-1/c

; Sequence 1, Application US/09557884

; Patent No. 6506581

; GENERAL INFORMATION:

; APPLICANT: Fleischmann et al.

; TITLE OF INVENTION: The Nucleotide sequence of

the Haemophilus influenzae Rd Genome, Fragments

Thereof, and Uses Thereof

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/557,884

FILING DATE: 25-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,102

FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB186P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs


```

;
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match      34.5%; Score 319.4; DB 3; Length 1830121;
Best Local Similarity 65.8%; Pred. No. 4.2e-98;
Matches 480; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

QY 85 AATTAGTAGCAACATCGACGCCGACATCTTATCGTTGAAGACGAGTTGGTAAAC 144
Db 936640 AGTAAAGATGAGAAATGACTACTCCAAATCTCGTTGTGAAGATAAAATTTGTCCAC 936581

QY 145 ACGCAACACGTTGAAAGATTTTCGAAGCGAAGCTATGATGTTTTCGAAGCGACAGA 204
Db 936580 TCGAATACGCTTAAGGATTTTGAAGCGAAGGATATGATGTTTGAAGCGAAGAA 936521

QY 205 TGGCGCGGAATGCAATCAGATCCTCTCTGAATATGACATCAACCTGGTGATCATGGATAT 264
Db 936520 TGGTGTGGAATGCATCATATATTTGGCAAAATCATATATTAATTTGGTTGTGATGATAT 936461

QY 265 CAATCTGCGGGTAAGAACGGTCTTCTGTTAGCGGTGAACCTGCGGACGAGGCGGAATGT 324
Db 936460 TAAATTTACCGGCAAAAACCGCTTATTATTGGCAAGAGAACTCCGTTGAAGAAATTAAGCTT 936401

QY 325 TGGCTTGATGTTCTGATCGGCGGTGACCAAGAGTTCGATAAAATTCGCGCTCGAAAT 384
Db 936400 ACCTCTTATTTTAACTGGTGGAGATAAAGTGGATATAAAATTTAGGTTTGGAAAT 936341

QY 385 CGGTGCAGATGACTACATCACCAACCGCTTCAACCCGCGGTGAACCTGACGATTCGTGCACG 444
Db 936340 TGGTCCGATGACTATTTGACTAAGCTTTTAAACCTAGAGAACTGACTATCCGCGCAG 936281

QY 445 CAACCTACTGTCGCGTACCATGAATCTGGGTACTGTGACGGAAGAACGTCGTAGGTTGA 504
Db 936280 TAAATTTGTCATCGTGCATG---CCGCATCAAGAAAAAGAAATACATTTGGTTCGAGA 936224

QY 505 AAGCTACAAAGTTCAATGTTGGGAATCGAGATCAACAGCGCTTCTGTTGATCGGCCCTGA 564
Db 936223 ATTCTATCGTTTAAATGTTGGAATTTAGACCTTAATAGCCACAGTTTAAATTAACCCAGA 936164

QY 565 TGGCGAGCAGTACAAGCTCGCGCGCAGCGAGTTCCGCGCCATGCTTCACTTCTGTGAAAA 624
Db 936163 AGGACAAGAAATCAAACTTCTCGCAGTGAATTTCTGTCATGTTTACATTTCTGTGAAA 936104

QY 625 CCCAGGCAAAATTCAGTCGCGGTGCTGAATCTGCTGAAGAAATGACCGCGCGTGAAGTGA 684
Db 936103 TCCAGGAAAAATTCGCAACCGCGCAAGAAATTTGCTGAAAAAATGACCGGACGAGATTAAA 936044

QY 685 ACCGACACCGTACTGTAGAGTCAAGTCCGCGGTATTCGTAACATTTTCGATCTAC 744
Db 936043 ACCTCAGGATCGTACCGTAGATGTACAAATTCGAGTATCGAAGAAACATTTTGAAGATCA 935984

QY 745 GCGCGATACGCGGAAATCATCGCCACCATTCACCGTGAAGGTATCGCTTCTGCGGTGA 804
Db 935983 TCCCAATFACTCAATATCATATGACTATGACTATGATGAGAGAGCTATCGTTTGTGGAGA 935924

QY 805 TCTGGAAGA 813
Db 935923 TATTGAATA 935915

RESULT 4
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
;

TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PBI86P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match      34.5%; Score 319.4; DB 3; Length 1830121;
Best Local Similarity 65.8%; Pred. No. 4.2e-98;
Matches 480; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

QY 85 AATTAGTAGCAACATCGACGCCGACATCTTATCGTTGAAGACGAGTTGGTAAAC 144
Db 936640 AGTAAAGATGAGAAATGACTACTCCAAATCTCGTTGTGAAGATAAAATTTGTCCAC 936581

QY 145 ACGCAACACGTTGAAAGATTTTCGAAGCGAAGCTATGATGTTTTCGAAGCGACAGA 204
Db 936580 TCGAATACGCTTAAGGATTTTGAAGCGAAGGATATGATGTTTGAAGCGAAGAA 936521

QY 205 TGGCGCGGAATGCAATCAGATCCTCTCTGAATATGACATCAACCTGGTGATCATGGATAT 264
Db 936520 TGGTGTGGAATGCATCATATATTTGGCAAAATCATATATTAATTTGGTTGTGATGATAT 936461

QY 265 CAATCTGCGGGTAAGAACGGTCTTCTGTTAGCGGTGAACCTGCGGACGAGGCGGAATGT 324
Db 936460 TAAATTTACCGGCAAAAACCGCTTATTATTGGCAAGAGAACTCCGTTGAAGAAATTAAGCTT 936401

QY 325 TGGCTTGATGTTCTGATCGGCGGTGACCAAGAGTTCGATAAAATTCGCGCTCGAAAT 384
Db 936400 ACCTCTTATTTTAACTGGTGGAGATAAAGTGGATATAAAATTTAGGTTTGGAAAT 936341

QY 385 CGGTGCAGATGACTACATCACCAACCGCTTCAACCCGCGGTGAACCTGACGATTCGTGCACG 444
Db 936340 TGGTCCGATGACTATTTGACTAAGCTTTTAAACCTAGAGAACTGACTATCCGCGCAG 936281

QY 445 CAACCTACTGTCGCGTACCATGAATCTGGGTACTGTGACGGAAGAACGTCGTAGGTTGA 504
Db 936280 TAAATTTGTCATCGTGCATG---CCGCATCAAGAAAAAGAAATACATTTGGTTCGAGA 936224

QY 505 AAGCTACAAAGTTCAATGTTGGGAATCGAGATCAACAGCGCTTCTGTTGATCGGCCCTGA 564
Db 936223 ATTCTATCGTTTAAATGTTGGAATTTAGACCTTAATAGCCACAGTTTAAATTAACCCAGA 936164

QY 565 TGGCGAGCAGTACAAGCTCGCGCGCAGCGAGTTCCGCGCCATGCTTCACTTCTGTGAAAA 624
Db 936163 AGGACAAGAAATCAAACTTCTCGCAGTGAATTTCTGTCATGTTTACATTTCTGTGAAA 936104

QY 625 CCCAGGCAAAATTCAGTCGCGGTGCTGAATCTGCTGAAGAAATGACCGCGCGTGAAGTGA 684
Db 936103 TCCAGGAAAAATTCGCAACCGCGCAAGAAATTTGCTGAAAAAATGACCGGACGAGATTAAA 936044

QY 685 ACCGACACCGTACTGTAGAGTCAAGTCCGCGGTATTCGTAACATTTTCGATCTAC 744
Db 936043 ACCTCAGGATCGTACCGTAGATGTACAAATTCGAGTATCGAAGAAACATTTTGAAGATCA 935984

QY 745 GCGCGATACGCGGAAATCATCGCCACCATTCACCGTGAAGGTATCGCTTCTGCGGTGA 804
Db 935983 TCCCAATFACTCAATATCATATGACTATGACTATGATGAGAGAGCTATCGTTTGTGGAGA 935924

QY 805 TCTGGAAGA 813
Db 935923 TATTGAATA 935915
```

```
QY 505 AAGCTCAAGTTCAATGTTGGAACTGGACATCAACAGCCGTTGTTGATCGGCCCTGA 564
Db 936223 ATTCTATCGTTTAAATGTTGAAATTAGACCTAAATAGCCACAGTTTAAATTACACCAGA 936164
QY 565 TGGCAGCAGTACAAGCTGCCGCGCAGCGAGTTCGCGCCCATGCTTCACTTCTGTGAAAA 624
Db 936163 AGGACAAGAAATTCAAACTTCCTCGCAGTGAATTTCTGTCGAATGTTTACATTTCTGTGAAAA 936104
QY 625 CCCAGGCAAAATTCAGTCCCGTGCTGAAGTGAAGAAATGACCGCGCGTGAGCTCAA 684
Db 936103 TCCAGAAAATTTGCAAAAGCGCGAAGAAATGCTGAAAAAATGACCGCAGAGAGTTAAA 936044
QY 685 ACCGCACGACCGTACTGTAGAGCTGACGATCGCGCGTATTCGTAACAATTTGCAATCTTAC 744
Db 936043 ACCTCAGGATCGTACCGTAGATGTCACAATTCGACGTATCAGAAAACATTTTGAAGATCA 935984
QY 745 GCCGATACGCCGGAATCATCGCCACCATTCAACGCTGAAGTTATCGTTCTCTGCGGTGA 804
Db 935983 TCCCAATATCTCAAAATATCATTTATGACTATATCATGGAAGGCTATCGTTTTTGTGGAGA 935924
QY 805 TCTGGAAGA 813
Db 935923 TATTGAATA 935915

RESULT 5
US-10-158-865-1/c
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; Patent No. 6846651
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2CID1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11924)..(11924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature

Query Match      34.5%; Score 319.4; DB 3; Length 1830121;
Best Local Similarity 65.8%; Pred. No. 4.2e-98;
Matches 480; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

QY 85 AATTTAGGTAGCAAAACATGACAGACCCCGCACATTTCTTATCGTTGGAAGACGAGTTGGTAAC 144
Db 936640 AGTATAAGGATGGAAATGACTACTCCAAAAATTTCTCGTTGTTGAAGATGAAATTTGTCCAC 936581

QY 145 ACGCAACACGTTGAAAAGTATTTTCGAAGCGGAAGGCTATGATGTTTTCGAAGCGACAGA 204
Db 936580 TCGAAATACGCTTAAAGGGATTTTTCGAAGCGGAAGGATATGATGTTTTCGAAGCGA 936521

QY 205 TGGCGCGGAAATGATCAGATCCTCTCTGAATATGACATCAACCTGGTGATCATCGATAT 264
Db 936520 TGGTGTGAAATGATCATATATTTGGCAATCATATATTAATTTTGGTTGTGATGATAT 936461

QY 265 CAATCTGCCGGTAAAGAACGGTCTTCTGTTAGCGGTGAATCTGCGCGAGCAGGCGGAATGT 324
Db 936460 TAATTTACCAGGCAAAAACGGCTTATTATTTCGCAAGAGNACTCCGTGAAGAAATTAAGCTT 936401

QY 325 TGGCTTGATGTTCTTGACTGSCCGTGACCAAGAGTGGATAAAATTTTCGGCGCTCGAAAT 384
Db 936400 ACCTCTTATTTTTTAACTGGTCGAGATAATGAAGATGGATAAAATTTTAGGTTTGGAAAT 936341

QY 385 CGGTGCAGATGACTACATCACAAACCGTTTCAACCCCGGTCACTGACGATTCGTGCGACG 444
Db 936340 TGGTGGCGATGACTATTGACTAAGCCTTTTAAACCCTAGAGAACTGACTATCCGCGCACG 936281

QY 445 CAACCTACTGTCCCGTACCATGAATCTGGGTACTGTCTCAGCGAAGAACGTCGTAGCGTTGA 504
Db 936280 TAATCTATTGCATCGTGAATG---CCGCATCAAGAAAAAGAAATACATTTTGGTCGAGA 936224

QY 505 AAGCTACAAGTTCAATGTTGGGAACTGGAGCATCAACAGCCGTTGTTGATCGGCGCCTGA 564
Db 936223 ATTCTATCGTTTAAATGGTTGGAAATTAGACCTAAATAGCCACAGTTTAAATTCACACCA 936164

QY 565 TGGCGAGCAGTACAGCTCCGCGCAGCGAGTTCGCGGCCACTGCTTCACTTCTGTGAAA 624
Db 936163 AGGACAAGATTCAAACTTCTCGCAGTGAATTTCTCGCAATGTTTACATTTCTGTGAAA 936104

QY 625 CCCAGGCAAAATTCAGTCCCGTGTGAACTGCTGAAGAAAAATGACCGGCCGTGAGCTGAA 684
Db 936103 TCCAGGAAATTTGCAACCGCGCGAAGAAATTCCTGAAAAAAATGACCGGACGAGAGTTAA 936044

QY 685 ACCGACGACCGTACTGTAGACGTGACGATCCGCGGCTATTTCGTAAACATTTTCGAATCTAC 744
Db 936043 ACCTCAGGATCGTACCGGTAGATGTCAAAATTCGACGTATCAGAAAAACATTTTGAAGATCA 935984

QY 745 GCCGATACGCGGGAATATCATCGCCACCATTCAGGTGAAGGTTATCGCTTCTCGGGTGA 804
Db 935983 TCCCAATACTCCAATATCATATGACTATATCATATGAGAGGCTATCGTTTTTGTGGAGA 935924

QY 805 TCTGGAAGA 813
Db 935923 TATTGAATA 935915
```

RESULT 6

```
US-09-252-991A-16417
; Sequence 16417, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
```

; TITLE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16417
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16417

Query Match 13.6%; Score 126; DB 3; Length 837;
Best Local Similarity 52.9%; Pred. No. 6.5e-33;
Matches 295; Conservative 0; Mismatches 260; Indels 3; Gaps 1;

QY 243 TCAACCTGGTGATCATGGATATCAATCTGCCGGGTAAAGACGGTCTTCTGTTAGCGCGTG 302
DB 263 TCAGAGCTGGTGCTGCTGGACATCCGCTGCCCGCAAGGACGGCTGACCTTGACCCGCG 322
QY 303 AACTGCGGAGCAGCGGAATGTTGGGTTGATGTTCTGACTGGCGGTGACCAACGAAGTCG 362
DB 323 AGCTGGGGTGGCTCCGAGGTGGGATCATCTGATCACGGGCGCAACGACGAGATCG 382
QY 363 ATAAATTTCTCGGCTCGAAATCGGTGCGAGATGATACATACCAAAACGGTTCAACCCGCG 422
DB 383 ATCGATCGTGGCGCTGGAGTGGCGCGCGACGATTAATGATCAAGCGCGTGAACCCGCG 442
QY 423 GTGAAGTACGATTCGTGCACCAACCTACTCTGCCGTACCATGAATCTGGGTACTGTCA 482
DB 443 GCGAATCTGTTGCGGGC---GAAGATCTGATCCGCGGTGGCCATGCCAGGCCA 499
QY 483 GCGAAGAACGTCGTAGCTTGAAGCTACAAGTTCAATGTTGGGAACCTGGACATCAACA 542
DB 500 GCGCGGCCCCCGCGGAGGCCCTCAGGACGTTCCGGGACCTGGCTGGAGCGCGACC 559
QY 543 GCGGTTGTTGATCGGCGCTGATGGGAGCAGTACAGTGGCGGCGAGCGAGTTCCGCG 602
DB 560 GCGCGCGCTGATCGACCAACGCGGCGCAACGAACCTGCTCACCCATGCGGAATTCGAAT 619
QY 603 CCATGCTTCACTTCTGTGAACCCAGGCAAAATTCAGTCCCGTCTGAATCTGCTGAAGA 662
DB 620 TCGTGGCGCTTCTTGGCGCAACAGCGGCATACCTGAGCGCGACCGAGTTGATGGACC 679
QY 663 AAATGACCGCGGTGAGCTGAACCGCACGACCGTACTGTAGACGTGACGATCCGCGCGTA 722
DB 680 AGATCGCAACCGCGAGTGGCTGCCAGCGATCGCTCCATCGATGCTGGTGGCGCGCC 739
QY 723 TTGCTAAACATTTGGAATCTAGCGCGGATACCGCGGAATCATCGCCACCATTCAGGTTG 782
DB 740 TGGCTGCAAGCTGCGCGACGACCGCGCGCAACCTGATCATCATCCATCCACGCGCG 799
QY 783 AAGGTTATCGTCTTCGCG 800
DB 800 CCGGCTACCTGTTTACCG 817

RESULT 7
US-09-252-991A-16171/c
; Sequence 16171, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16171
; LENGTH: 2430
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16171

Query Match 13.6%; Score 126; DB 3; Length 2430;
Best Local Similarity 52.9%; Pred. No. 1.3e-32;
Matches 295; Conservative 0; Mismatches 260; Indels 3; Gaps 1;

QY 243 TCAACCTGGTGATCATGGATATCAATCTGCCGGGTAAAGACGGTCTTCTGTTAGCGCGTG 302
DB 1406 TCGAGCTGGTGCTGCTGGACATCCGCTGCCCGCAAGGACGGCTGACCTTGACCCGCG 1347
QY 303 AACTGCGGAGCAGCGGAATGTTGGGTTGATGTTTCTGACTGGCGGTGACCAACGAAGTCG 362
DB 1346 AGCTGGGGTGGCTCCGAGGTGGGATCATCTGATCACGGGCGCAACGACGAGATCG 1287
QY 363 ATAAATTTCTCGGCTCGAAATCGGTGCGAGATGATACATACCAAAACGGTTCAACCCGCG 422
DB 1286 ATCGATCGTGGCGCTGGAGTGGCGCGCGCGACGATTACGTGATCAAGCGCGTGAACCCGCG 1227
QY 423 GTGAAGTACGATTCGTGCACGCAACCTACTCTGCCGTACCATGAATCTGGGTACTGTCA 482
DB 1226 GCGAATCTGTTGCGGGC---GAAGATCTGATCCGCGGTGGCCATGCCAGGCCA 1170
QY 483 GCGAAGAACGTCGTAGCGTTGAAGCTACAAGTTCAATGTTGGGAACCTGGACATCAACA 542
DB 1169 GCGCGGCCCCCGCGGAGGCCCTCAGGACGTTTCGGGACTGCTGCTGGACGCGGACC 1110
QY 543 GCGGTTGTTGATCGGCGCTGATGGCGAGCAGTACAGTCCGCGGCGAGCGAGTTCCGCG 602
DB 1109 GCGCGCGCTGATCGACCAACGCGGCGCAACGAACCTGCTCACCCATGCGGAATTCGAAT 1050
QY 603 CCATGCTTCACTTCTGTGAACCCAGGCAAAATTCAGTCCCGTCTGAACTGCTGAAGA 662
DB 1049 TGCTGGCGCTTCTTGGCGCAACAGCGGCGCATACCTGAGCGCGGACGAGTTGATGGACC 990
QY 663 AAATGACCGCGGTGAGCTGAACCGCACGACCGTACTGTAGACGTGACGATCCGCGCGTA 722
DB 989 AGATCGCAACCGCGAGTGGCTGCCAGCGATCGCTCCATCGATGCTGGTGGCGCGCC 930
QY 723 TTGCTAAACATTTGGAATCTAGCGCGGATACCGCGGAATCATCGCCACCATTCAGGTTG 782
DB 929 TGGCTGCAAGCTGCGCGACGACCGCGCGCAACCGCGCAACTGATCATCATCCATCCACGCGG 870
QY 783 AAGGTTATCGTCTTCGCG 800
DB 869 CCGGCTACCTGTTTACCG 852

RESULT 8
US-09-252-991A-16309
; Sequence 16309, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16309
; LENGTH: 3675
; TYPE: DNA

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16309

Query Match 13.6%; Score 126; DB 3; Length 3675;
Best Local Similarity 52.9%; Pred. No. 1.7e-32;
Matches 295; Conservative 0; Mismatches 260; Indels 3; Gaps 1;

QY 243 TCAACCTGGTATCATGGATATCAATCTGCGGGTAAGAACGGTCTTCTGTTAGCGCGTG 302
DB 72 TCGAGCTGGTGTCTGGACATCCGCTGCGCGCAAGAGCGGCTGACCCCTGACCCGCG 131

QY 303 AACTGCGGAGCAGCGGAATGTTGGTTGATGTTCTGACTGCGCGTGAACAGAGTCG 362
DB 132 AGCTCGGGTGGCTCCGAGGTGGGGATCATCTGATCACCGGGGCAACGACGAGATCG 191

QY 363 ATAAAAATCTCGGCTCGAAATCGGTGAGATGACTACATACCAAAACGGTTCAACCCGC 422
DB 192 ATCGCATCGTGGCTGAGTGGCGCGGACGATTAATGATCAAGCGCTGAACCCGC 251

QY 423 GTGAATGACGATTCGTGCAGCAACCTACTGTCCCGTACCATGAATCTGGGTACTGTCA 482
DB 252 GCGAACTGTGTGCGGGC---GAAGAATCTGATCGCGGGTGGCCATGCCAGGCCA 308

QY 483 GCGAAGAGCTGTAGCTGTTGAAGCTCAAGTTCAATGTTGGGAACTGGACATCAACA 542
DB 309 GCGCGGGCCCGCGGAGGCGCTCAGCGAGTTCGGGAGCTGGCTGTGGACGCGGACC 368

QY 543 GCGCTTGTGTATCGGCTGATGCGGAGCAGTACAAGCTGCGCGCAGCGAGTTCCGCG 602
DB 369 GCGCGGCTGTATGACACACGCGGGCAAGAAACCTGCTCACCATGGCGAATCCAAAT 428

QY 603 CCATGCTTCACTTCTGTGAACCCAGGCAAAATTCAGTCCCGTCTGAACTGTCTGAAGA 662
DB 429 TGCTCGGCGCTTCTCGGCAACAGCGGCATACCTTGAGCGCGACAGTTGATGGACC 488

QY 663 AAATGACCGCGGTGAGCTGAACCGCAGCAGCTGATGTAGAGTGAAGTTCGCGGTA 722
DB 489 AGATCGCAACCGAGTGGCTGCGCGAGCTGCTCCATCGATGCTGTGGTGGCGCGCC 548

QY 723 TTCTGAAATTTTGAATCTACGCGGATACGCGGAATCATGCCACCATTCACGCTG 782
DB 549 TCGCTGCGAGTGGCGGAGCAGCCCGCGAACCAGTGAATCATCCATCCACGCGG 608

QY 783 AAGTTATCGTTCTGCG 800
DB 609 CCGGTACCTGTTCCCG 626

RESULT 9

US-09-252-991A-11517/c
; Sequence 11517, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 11517

LENGTH: 990

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11517

Query Match 13.0%; Score 120.6; DB 3; Length 990;
Best Local Similarity 49.2%; Pred. No. 5.6e-31;
Matches 349; Conservative 0; Mismatches 354; Indels 6; Gaps 1;

QY 95 GCAAAACATGCAAGCCCGCAATCTTATCTGTTGAAGACGAGTTGGTAACACGCAACACG 154
DB 747 GCAATGGAGCATGTCGATCACAATCTGATCTGTCAGCATGACCGCGAGATTCGCGAGCTG 688

QY 155 TTGAAAAGTATTTTCGAAGCGGAAGGCTATGATGTTTTCGAAGCGACAGATGGCGCGGAA 214
DB 687 GTAGGCAACTACTTGAAGAAGACGGCTGCGCACCAACATCATCTGTCGCGCATGGCGCGAG 628

QY 215 ATGCATCAGATCTCTCTGAATATGACATCAACCTGGTGATCATGATATCAATCTGCGG 274
DB 627 ATGCGGGTTCCTCGAAGCCACAACGGTGAACCTGATCTGCTCTCGAATCATGATGCC 568

QY 275 GGTAAAGACGGTCTTCTGTTAGCGGTGAACCTGGCGG-----AGCAGCGCAATTTGCG 328
DB 567 GCGCAGCAGCGCTGCTCTGCGCGAGCTGCGGGTGGCGGCAAGCAAGCCACGCGG 508

QY 329 TTGATGTTCTGACTGCGCGGTGACAAACGAAGTGAATAAATTTCTCGGCTCGAAATCGGT 388
DB 507 GTGCTGATGCTTCAACCGCGCGCAACGAGACCGACCGCATCATCGGCTGGAGATGGGC 448

QY 389 GCAGATGACTACATCACCACCGTTCAACCGCGTGAACCTGACGATTCGTGACGCAAC 448
DB 447 GCGCAGCACTACCTGACCAAGCCCTTCTCCGCGCGGAACCTGTTGCCCGGATCAACGCC 388

QY 449 CTACTGTCCTCCATCAATCTGAGTACTGTACGCGAAGAACTGCTAGCGTTTGAAGC 508
DB 387 GTGCTGCGGCTACCGCGATGCTCCGCGCAACCTCACCGTGAGCGAGCAGCGCGCTG 328

QY 509 TACAAGTTCAATGTTGGAACTGACATCAACAGCGGTTGTTGATCGGCGCTGATGGC 568
DB 327 ATCGGCTTCGGCCAGTGGCAACTGACACACGAGCGCGCCACCTGCTCGACGACCGCGC 268

QY 569 GAGCAGTACAAGCTGCGCGCAGCGAGTTCCGCGCAATGCTTCACTTCTGTGAACCCCA 628
DB 267 ACCGTGTCGCGCTCAGCGCGCGAGTACCGCTGTCGGGTTTCTCGATCATCCG 208

QY 629 GGCATAATTCAGTCCCGTGTGAATGCTGAAGAAATGACCGCGCGTGAAGTGAACCG 688
DB 207 CAACGGGTACTCAGCGCGGACAGTTGCTCAACCTGACCCAGGGCGCGAGCGGACATC 148

QY 689 CACGACCGTACTGATGAGTGAAGTACGATCCGCGTATTGTTAAACATTTTGAATCTAGCGG 748
DB 147 TTCGACCGTTTCCATCGACCTGCTGTCAGCGCGCTGCGCCCAACGCTCGCGACGACGCC 88

QY 749 GATACGCGGAAATCATCGCCACCATTCACGCTGAAGTTATCGCTTCT 797
DB 87 CGCGAACCGGAGTACATCAAGACCGTGGCAGCGGGCTATGTGTTCT 39

RESULT 10

US-09-252-991A-11722
; Sequence 11722, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 11722

LENGTH: 2166

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11722

Query Match 13.0%; Score 120.6; DB 3; Length 2166;

[illegible]

2083 CGGAGCCGAGGATACATCAGACGCGTGGCGAGCGAGGCGCTAIGTGTCT 2131
 DD
 RESULT 11
 US-09-252-991A-11865
 ; Sequence 11865, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252.991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 11865
 ; LENGTH: 2004
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-11865

Query Match	12.3%	Score 113.8	DB 3	Length 2004
Best Local Similarity	50.6%	Pred. No. 2.1e-28		
Matches	305	Conservative	0	Mismatches 292
			Indels	6
			Gaps	1

QY	201	CAGATGCGCGGAAATGCATCAGATCTCTCTGAATATGACATCAACTCGTGTGATCATGG	260
DB	66	CCGATGCCCGCAGATCGCGGCTTCTCGAAGCAACACGGTGGACCTGATCGTCTCTCG	125
QY	261	ATATCAATCTCCCGGGTAAGAACGGTCTTCTGTGTAGCGCGGTGAATCGCG-----AGC	314
DB	126	ACATCATGATGCCCGCGGACGACGGCTGTCTCTCGCGGAGCTGCGCGTCGCGCAAGC	185
QY	315	AGGCGAATGTTGGTGTGATGTTCTGTAGCTGGCGGTGACACGAGTGCATATAAATTCCTCG	374
DB	186	ACAAGGCCACGCCGGTGTCTGATGTCTACCGCGCGCAACGACGACCGCATCATCG	245
QY	375	GCCTCGAAATCGGTGCAGATGACTATACATACCAAAACGGTTCAAACCGCGTGAATGACGA	434
DB	246	GCCTGGAGATGGGCGCGACGACTACTCTGACCAAGCCCTTCTCGCGCGGACCTGCTTG	305
QY	435	TTCTGTCAACGCAACTACTGTCCCGTACCATGAATCTGGGTACTGTGACGCAAGAACGTC	494
DB	306	CCCGGATCAACGCCGTGTGCGGGTACGCGGATGCTGCCCGCAACCTCACCGGTGAGCG	365
QY	495	GTACGCTTGAAGCTACAGTTCAATGGTTGGGAACCTGGACATCAACGACCGTTCGTTGA	554
DB	366	AGAGCAGCCCGCTGATCGGCTTCGGCTTCGGCGAGTGGCACTGGACACGAGCGCCGCA	425
QY	555	TCGCGCCCTGATGGGAGCAGTACAAAGCTGCGCGCAGCGAGTTCGCGGCCATGCTTCACT	614
DB	426	TCGACGAGCGCGGACCGTGTGCTGCCCTCAGCGCGCGGAGTAGCGCTGCTCGGGTGT	485
QY	615	TCTGTGAAAACCCAGGCAAAATTTCAGTCCCGTGTCTGAACCTGCTGAAGAAAATGACCGCC	674
DB	486	TCCTCGATCATCGCAACGGGTACTCAGCGCGCAGCACAGTGTCTCAAACCTGACCCAGG	545
QY	675	GTGAGCTGAAACCCGACGACCGTACTGTAGACGTGACGATCCGCGGTATTCGTAACACT	734
DB	546	GCGAGGCGGACATCTTGACCGTTCCATGCACTGCTGGTTCAGCGCGCTGCGCCACGCC	605
QY	735	TCGAATCTAGCCCGGATACGCCGGAATCATCGCCACCATTCAGGTGAAGGTTATCGCT	794
DB	606	TCGCGGACGACCGCCGCGAACCAGGATACATCAAGACCGTGCAGCGAGGGCTATGTGT	665
QY	795	TCT 797	
DB	666	TCT 668	


```

RESULT 12
US-09-418-980-1
; Sequence 1, Application US/09418980
; Patent No. 6713071
; GENERAL INFORMATION:
; APPLICANT: Campos, Manuel
; APPLICANT: Baarsch, Mary Jo
; APPLICANT: Rosey, Everett
; APPLICANT: Ankenbauer, Robert
; APPLICANT: Warren-Stewart, Lynn
; APPLICANT: Suiter, Brian
; APPLICANT: Keach, Robin
; TITLE OF INVENTION: NOVEL PROTEINS FROM ACTINOBACILLUS PLEUROPNEUMONIAE
; FILE REFERENCE: PC9854A
; CURRENT APPLICATION NUMBER: US/09/418,980
; CURRENT FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1018
; TYPE: DNA
; ORGANISM: Actinobacillus pleuropneumoniae
; FEATURE:

```

```
; NAME/KEY: CDS
; LOCATION: (272)..(787)
US-09-418-980-1

Query Match      11.6%; Score 107.4; DB 3; Length 1018;
Best Local Similarity 80.3%; Pred. No. 2.3e-26;
Matches 126; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 662 AAAATGACCGCGGTGAGCTGAACCGCAGCAGCGTACTGTAGAGCTGACGATCGCGGT 721
Db 68 AAAATGACCGCGGTGAGCTGAACCGCAGCAGCGTACTGTGATGACGATTCGTGT 127

Qy 722 ATTGCTAAACATTTTGAATCTACGCGCGGATACGCGGAAATCATCGCCACCATTCACGGT 781
Db 128 ATTGCTAAACATTTTGAAGATCACCTAATACACCGGAAATCATTTGTAACCATTCATGGT 187

Qy 782 GAAGGTTATCGTTCTGCGGTGATCTGGAAGTTAAT 818
Db 188 GAAGGTTATCGTTTTCGCGGAGTTAGAGTAGTAAT 224

RESULT 13
US-09-634-238-133/C
; Sequence 133, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christensen, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; FILE REFERENCE: 11000.1043U1
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2006-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-133

Query Match      10.6%; Score 98.6; DB 3; Length 729;
Best Local Similarity 46.3%; Pred. No. 2.1e-23;
Matches 323; Conservative 0; Mismatches 374; Indels 0; Gaps 0;

Qy 156 TGAAGAATTTTGAAGCGGAGGCTATGATGTTTTCGAAGCGACAGATGGCGGAAA 215
Db 728 TCAAGTTAAATTTGGATATAAGAGGCTATGTCCTCACTCATATGATGCGGAAG 669

Qy 216 TGCATCAGATCTCTCTGAATATGACATCAACCTGGTATCATGATATCAATCTGCCGG 275
Db 668 CTTGAAGAAGTTGAATCAGAAATCACCTGATTTGATTTACTTGAATGATGTACC 609

Qy 276 GTAAGAAGGTTCTTCTGTAGCGCTGAATCGCGGAGCAGCGGAATGTTGCGTTGATGT 335
Db 608 AAATGATGGGCTTGAAGTTGCCCGACAAATTCGTAAGACATGACACGCGGATTA 549

Qy 336 TCTGATGGCGGTGACACGAAGTCGATAAAATTCCTGGGCTCGAAATCGGTGAGATG 395
Db 548 TGCTGACGCGGAAGGATTCGGAATTTGATAAAGTGTAGGACTGAGCTTGGGGCAGATG 489

Qy 396 ACTACATACCAACCGTTCAACCGCGGTGACATGACGATTCGTGACGCAACCTACTGT 455
Db 488 ATTATGTTACCAAGCGGTTTTCCTCAATCAGAACTTTGTCGCGGAGTGAAGGCTA 429

; NAME/KEY: CDS
; LOCATION: (272)..(787)
US-09-418-980-1

Query Match      11.6%; Score 107.4; DB 3; Length 1018;
Best Local Similarity 80.3%; Pred. No. 2.3e-26;
Matches 126; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 662 AAAATGACCGCGGTGAGCTGAACCGCAGCAGCGTACTGTAGAGCTGACGATCGCGGT 721
Db 68 AAAATGACCGCGGTGAGCTGAACCGCAGCAGCGTACTGTGATGACGATTCGTGT 127

Qy 722 ATTGCTAAACATTTTGAATCTACGCGCGGATACGCGGAAATCATCGCCACCATTCACGGT 781
Db 128 ATTGCTAAACATTTTGAAGATCACCTAATACACCGGAAATCATTTGTAACCATTCATGGT 187

Qy 782 GAAGGTTATCGTTCTGCGGTGATCTGGAAGTTAAT 818
Db 188 GAAGGTTATCGTTTTCGCGGAGTTAGAGTAGTAAT 224

RESULT 13
US-09-634-238-133/C
; Sequence 133, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christensen, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; FILE REFERENCE: 11000.1043U1
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2006-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-133

Query Match      10.6%; Score 98.6; DB 3; Length 729;
Best Local Similarity 46.3%; Pred. No. 2.1e-23;
Matches 323; Conservative 0; Mismatches 374; Indels 0; Gaps 0;

Qy 156 TGAAGAATTTTGAAGCGGAGGCTATGATGTTTTCGAAGCGACAGATGGCGGAAA 215
Db 728 TCAAGTTAAATTTGGATATAAGAGGCTATGTCCTCACTCATATGATGCGGAAG 669

Qy 216 TGCATCAGATCTCTCTGAATATGACATCAACCTGGTATCATGATATCAATCTGCCGG 275
Db 668 CTTGAAGAAGTTGAATCAGAAATCACCTGATTTGATTTACTTGAATGATGTACC 609

Qy 276 GTAAGAAGGTTCTTCTGTAGCGCTGAATCGCGGAGCAGCGGAATGTTGCGTTGATGT 335
Db 608 AAATGATGGGCTTGAAGTTGCCCGACAAATTCGTAAGACATGACACGCGGATTA 549

Qy 336 TCTGATGGCGGTGACACGAAGTCGATAAAATTCCTGGGCTCGAAATCGGTGAGATG 395
Db 548 TGCTGACGCGGAAGGATTCGGAATTTGATAAAGTGTAGGACTGAGCTTGGGGCAGATG 489

Qy 396 ACTACATACCAACCGTTCAACCGCGGTGACATGACGATTCGTGACGCAACCTACTGT 455
Db 488 ATTATGTTACCAAGCGGTTTTCCTCAATCAGAACTTTGTCGCGGAGTGAAGGCTA 429
```

```
Qy 456 CCCGTACCATGAATCTGGGTACTGTCTCAGCGAAGAACGTCGTAGCTTTGAAAGCTACAAGT 515
Db 428 GCCGTACTCATCAGCCAAATGCTGCCGCGAGTGAGGAAGACCAAGCCAAATCGTGAATGG 369

Qy 516 TCATGTTGGGAAGTGGACATCAACAGCCGTTTGGTTGATCGGCCCTGATCGCGAGCAGT 575
Db 368 AAGTTGGCGAATTTAAACGATTCATCCGAGCGGTATACGGTATCTAAACGCGGTGAGAATA 309

Qy 576 ACAAGCTCCGCGCAGCGAGTTCGCGCCATGCTTCACTTCTGTGAAACCCAGGCAAAA 635
Db 308 TCGAATTAACCCACCGGAGTTTGAATTTGCTGCAATACCTTGGCCGCCACTTGGGACAGG 249

Qy 636 TTCAAGTCCGCGTGAATCTGCTGAAGAAATGACCGCGCGTGAAGTGAACCGCACGACC 695
Db 248 TTATGACCCGCGAACAATCTGCTACAGACAGTTTGGGGCTATGACTACTTTGGCGATGTCC 189

Qy 696 GTACTGTAGAGCTGACGATCCGCGTATTCGTAACATTTGCAATCTACGCGCGGATAGCG 755
Db 188 GGACTGTGGATGTGACGGTTGCGCCCTCGCGGAGAGATCGAAGACAATCCGTCGCCATC 129

Qy 756 CGGAATCATCGCCACCAATTCACGGTGAAGGTTATCGCTTCTGCGGTGATCTGGAAGATT 815
Db 128 CTGAGTGGTTGGTCAACCGGCGCGGTAGGCTATTTTGGGAATCCTGACGCTGAAT 69

Qy 816 AATCGGCTTTACCAACCGTCAAAAAAAGCGCGCTTTT 852
Db 68 AATTAGGAGCACCACTGTTGAATAAAAAAGATTGCTT 32

RESULT 14
US-09-252-991A-6307
; Sequence 6307, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6307
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6307

Query Match      10.5%; Score 97.2; DB 3; Length 762;
Best Local Similarity 49.4%; Pred. No. 6.7e-23;
Matches 339; Conservative 0; Mismatches 338; Indels 9; Gaps 3;

Qy 114 ACATTCTTATCGTTGAAGACGAGTTGGTAAACACGCAACACCGTTGAAAAGTATTTTCGAAG 173
Db 53 AGATCCTCGTCTGACGACGACGCGCCGCTGCGCGCTGCTCGAACGTTTCTCTCGACG 112

Qy 174 CGGAAGGCTATGATGTTTTTGAAGCGACAGATGCGCGGGAATGATGATCAGATCTCTCTG 233
Db 113 AGCAGGGCTACCGGCTCGCGCGGTGGAGAACACCGAGCAGATGATCGCTTCTATCCC 172

Qy 234 AATATGACATCAACCTCGTGTATCATGATATCAATTCGCGGGTAAGAACGCTTCTGT 293
Db 173 GCGAACTGTTCCAGTTGGTTGCTTCGACCTGATGCTTCCCGGAGACGCGCTCTACTG 232

Qy 294 TAGCGCGTGAACCTCGCGGAGCAGCGGAAT---GTTGCGTTGATGTTCTTGACTTGGCGGTG 350
Db 233 CTTGCCCGCGCTCGCGGAACAGAACAAATCAGGTGCGGATCATCATGCTCACCGCAAGG 292

Qy 351 ACAACGAAGTCGATAAAATTCGCGCTCGAAATCGGTCGAGATGACTATCATCATCAACAAC 410
Db 351 ACAACGAAGTCGATAAAATTCGCGCTCGAAATCGGTCGAGATGACTATCATCATCAACAAC 410
```

Db 293 GCGACGAGCGGACCGCATCCAGGCGCTGGAACCTGGGCGCGACGACTACCTGGCCAAAGC 352
Qy |||||
Db 411 CGTTCAACCCCGTGAATCGTACGATTCTGTGCACGCAACTACTACTGTCCGTACCATGAATC 470
Qy |||||
Db 353 CGTTCAACCCCGCGAACTGTCTGGCGCGGATCAAGGCCGTGTCTGGCGCGCCAGGCGCGCGC 412
Qy |||||
Db 471 TGGGTACTGTGACGGAAGAACGTCGTAGCGTTGTTGAAAGCTACAAAGTTCAATGGTTGGGAAC 530
Qy |||||
Db 413 TGGTCC---CGGCGCGCTGCGGTGCGGACGAGGTGTGACCTTCGGGCACTACCGAC 469
Qy |||||
Db 531 TGGACATCAACAGCCGTTGTTGATCGGCCCTTGATGGCGAGCAGTACAAAGCTGCGCGCA 590
Qy |||||
Db 470 TGTTCCTCGCCACCCGCGAACTGAAG---AAGGGCGACGAGTACACATGCTGACCAACCG 526
Qy |||||
Db 591 GCGAGTTCCGCCCATGCTTCACTTCTGTGAAACCCAGGCAAAATTCAGTCCCGTGTG 650
Qy |||||
Db 527 GCGAGTTCCCGTGTCAAGGCCCTGGTGCAGCACGCCCGCGAAACCGCTGACCCGGGACA 586
Qy |||||
Db 651 AACTGCTGAAGAAATGACCGCGCTGAGCTGAAACCGCACGCGTACTGTAGACGTGA 710
Qy |||||
Db 587 AGCTGATGAACCTCGCCCGTGGCCGCGAGTGGGATGCCCTGGAGGTTCCATCGACGTGC 646
Qy |||||
Db 711 CGATCCGCGGTATTCGTAAACATTTGCAATCTACCGCGGATACGCGCGAAATTCATCGCCA 770
Qy |||||
Db 647 AGATCTCGCGCTGCGCGCTGATCGAGCGGATCCGTCCAAGCGCGCTATATCCAGA 706
Qy |||||
Db 771 CCAATTCAGGTGAAGGTTATCGCTTC 796
Qy |||||
Db 707 CGGTCTGGGGCGTGGCTACGTGTTTC 732

RESULT 15

US-09-134-001C-226
; Sequence 226, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 226
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-226

Query Match 10.4%; Score 96.2; DB 3; Length 705;
Best Local Similarity 48.5%; Pred. No. 1.4e-22;
Matches 330; Conservative 0; Mismatches 343; Indels 8; Gaps 2;

Qy 117 TTCCTATCGTTGAAGACGAGTTGGTAAACGCAACACGTTGAAAGTATTTTCGAAGCGG 176
Db |||||
Qy 177 AAGGCTATGATGTTTTCGAAGCGACAGATGGCGCGGAAATGCATCAGATCCTCTCTGAAT 236
Db |||||
Qy 77 AAGGTTACGACGTATATTGCGCTTATGCGGTAATGACGCGTAGATTTAATCTATGAAG 136
Db |||||
Qy 237 ATGACATCAACCTGTGTGATCATGATATCAATCTGCGGGTGAAGAACGGTCTTCTGTAG 296
Db |||||
Qy 137 AAGAACCAGATATCGTCTTACTTGTATATCATGTACCTGGTAGAGATGTTATGGAAGTAT 196
Db |||||
Qy 297 CCGGTGAACCTGCGGACGAGGCGAATGTTGCGTTGATGTTCTGACTGCGCGTGACAACG 356
Db |||||
Qy 197 GTCGTGAAGTGGTAAAAAGTAAATGCGGATTCATGCTGACGCGAAAGATTCTG 256

Search completed: February 13, 2006, 14:41:47
Job time : 219 secs

Qy 357 AAGTCGATAAAATCTCGGCTCCGAATCGGTGCAGATGACTACATCACCACAAACCGTTCA 416
Db |||||
Qy 257 AAATTGATTAAGTATTAGGCTTGAATTTAGGTGCAGATGATTACGTAACCAATTTA 316
Qy 417 ACCGCGTGAACTGACGATTCGTGCACGCAACCTACTGTCCCGTACCAATGAATCTGGGTA 476
Db |||||
Qy 317 GTACTCGTGAACCTCATCGCACGTG-----TGAAGCGAACTTTACGCCGTCATTATTCA 369
Qy 477 CTGTTCAGGAAGAACGTCGTAGCCGTTGAAAGCTACAAAGTTCAATGGTTGGGAACCTGGACA 536
Db |||||
Qy 370 CAACGAGCTCAAGAAAGTAAGTGGTGCACAAATGAAATTTACAAATTAAGATATTGTGATT 429
Qy 537 TCAACAGCGG--TTGTTGATCGGCCCTGATGGCGAGCAGTACAAAGCTGCCGCGACGCGAG 595
Db |||||
Qy 430 TATCCAGATGCATATTCAATTTAAAAACGTTGGAGAGACATTGAAATTAACGATCGTGAA 489
Qy 596 TTCCGCGCCATGCTTCACTTCTGTGAAACCCAGGCAAAATTCAGTCCCGTGGTGAACGTG 655
Db |||||
Qy 490 TTCGAGCTATTTCATTTCTTAAACATATGGGTCAAGTCATGACACGCTGAACACTTA 549
Qy 556 CTGAAGAAATGACCGGCCGTGAGCTGAAACCGCACGACCGTACTGTAGAGTGAAGATC 715
Db |||||
Qy 550 CTACAAACAGTGTGGGGTTACGATTATTTTGGTGTATGTTCTGTGACGCTAACAAAT 609
Qy 716 GCGGTATTTCGTAAACATTTTCGAATCTACGCGCGATACGCGGAAATCATCGCCACCAAT 775
Db |||||
Qy 610 CGCGTTTAAAGAGAAATAATTTGAAGATGATCCATCTCATCCAGAAATACATTTGTGACAGT 669
Qy 776 CACGTTGAAGGTTATCGCTTC 796
Db |||||
Qy 670 AGAGCGTTGGATACTTCTC 690

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2006, 13:54:39 ; Search time 834 Seconds
(without alignments)
9191.503 Million cell updates/sec

Title: US-10-613-990A-31
Perfect score: 927
Sequence: 1 gtcattgtacgcccgcgatg.....gtactttcttactggcttt 927

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main: *
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq: *
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq: *
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq: *
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq: *
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq: *
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq: *
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq: *
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq: *
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	927	100.0	927	8 US-10-613-990A-31	Sequence 31, Appl
2	925.8	99.9	1732	6 US-10-264-237-983	Sequence 983, Appl
3	508.6	54.9	759	8 US-10-613-990A-19	Sequence 19, Appl
4	447.4	48.3	1048	9 US-10-450-763-28743	Sequence 28743, A
5	320.6	34.6	191996	9 US-10-795-159-683	Sequence 683, Appl
6	319.4	34.5	1830121	7 US-10-329-670-1	Sequence 1, Appli
7	319.4	34.5	1830121	8 US-10-158-865-1	Sequence 1, Appli
8	319.4	34.5	1830121	9 US-10-981-687-1	Sequence 1, Appli
9	316.2	34.1	711	3 US-09-815-242-7046	Sequence 7046, Ap
10	316.2	34.1	711	7 US-10-282-122A-22157	Sequence 22157, A
11	232.8	25.1	1110	9 US-10-450-763-18226	Sequence 18226, A
12	161.6	17.4	2125	9 US-10-795-159-291	Sequence 291, Appl
13	153.6	16.6	2227	9 US-10-795-159-230	Sequence 230, Appl
14	107.4	11.5	1018	8 US-10-770-824-1	Sequence 1, Appli
15	106.8	11.5	11275	6 US-10-240-689-39	Sequence 39, Appl
16	103.8	11.2	714	7 US-10-282-122A-35444	Sequence 35444, A
17	101.4	10.9	705	7 US-10-282-122A-35988	Sequence 35988, A
18	99	10.7	732	7 US-10-282-122A-14497	Sequence 14497, A
19	98	10.6	714	7 US-10-282-122A-24570	Sequence 24570, A
20	96.8	10.4	2256646	7 US-10-470-555-1	Sequence 1, Appli
21	96.2	10.4	705	7 US-10-724-972A-3136	Sequence 3136, Ap
22	96.2	10.4	721	7 US-10-282-122A-34637	Sequence 34637, A
23	95	10.2	1163020	7 US-10-398-221-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-10-613-990A-31
; Sequence 31, Application US/10613990A
; Publication No. US20040180404A1
; GENERAL INFORMATION:
; APPLICANT: Ajinomoto Co., Inc.
; TITLE OF INVENTION: Method for Producing Target Substance by Fermentation
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/613,990A
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: JP 2002-203764
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 927
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(817)
US-10-613-990A-31

Query Match	100.0%	Score 927;	DB 8;	Length 927;
Best Local Similarity	100.0%	Pred. No. 2.1e-300;	Mismatches 0;	Indels 0; Gaps 0;
Matches 927;	Conservative 0;			
QY	1	GTCATGTTACGCCGATCATGTTAAATTTTCAGCATGTCATCAGCGAGTTCAGGACTTTTGT	60	
Db	1	GTCATGTTACGCCGATCATGTTAAATTTTCAGCATGTCATCAGCGAGTTCAGGACTTTTGT	60	
QY	61	ACTTCTCTGTTTCGATTTAGTTGGCAATTTAGTAGCAAAACATGCAAGCCCGCACATTTCT	120	
Db	61	ACTTCTCTGTTTCGATTTAGTTGGCAATTTAGTAGCAAAACATGCAAGCCCGCACATTTCT	120	
QY	121	TATCGTTGAACAGCAGTTGGTAAACACGCAACACGTTTGAAGAAAGTATTTTCGAAGCGGAAAGG	180	
Db	121	TATCGTTGAACAGCAGTTGGTAAACACGCAACACGTTTGAAGAAAGTATTTTCGAAGCGGAAAGG	180	
QY	181	CTATGATGTTTTTCGAAGCGCAGATGCGCGGAAATGCATCAGATCCCTCTCTGAATATGA	240	
Db	181	CTATGATGTTTTTCGAAGCGCAGATGCGCGGAAATGCATCAGATCCCTCTCTGAATATGA	240	
QY	241	CATCAACCTGTTGATCATGGATATCAATCTGCCCGGTGAAGACGGTCTTCTGTTAGCGCG	300	
Db	241	CATCAACCTGTTGATCATGGATATCAATCTGCCCGGTGAAGACGGTCTTCTGTTAGCGCG	300	

QY	301	TGAATCGCGCAGCAGGCGGAATGTTGCGTTGATGTTCTGACTGGCGCGTGCAACGAACT	360
DB	301		
QY	361	CGATAAATTTCTCGCGCTCGAAATCGGTGCAGATGACTACATCACCAAACCGTTCACACC	420
DB	361		
QY	421	GGGTGAATCGACGATTCGTGACGCAACCTACTGTCCGTACCAATGATCTGGGTACTGT	480
DB	421		
QY	481	CAGCGAAGAACGTCGTAGCGTTGAAAGCTCAAAGTTCAAATGGTTGGGAATCGACATCAA	540
DB	481		
QY	541	CAGCGGTTCGTTGATTCGGCCCTGATGGCGAGCAGTACAAGTTCGGCGCGAGCGAGTTCGG	600
DB	541		
QY	601	CGCCATGCTTCACTTCTGTGAAAACCCAGGCAAAAATTCAGTCCGTGCTGAACTGCTGAA	660
DB	601		
QY	661	GAAATGACCGCGCGTGAGCTGAAACCGCAGCGTACTGTGTAGACGTGACGATCCGCCG	720
DB	661		
QY	721	TATTCGTAAACATTTTCGAATCTACGCCGGATACGCCGGAAATCATCCGACCAATTCACCG	780
DB	721		
QY	781	TGAAGTTATTCGCTTCGCGTGATCTCGAAGATTAATTCGGCTTTACCCCGTCAAAAA	840
DB	781		
QY	841	AACGGCGCTTTTACGGCGGTTTTATTTTTTCAACCTTTATTTCCAGATAGTAACCTCATC	900
DB	841		
QY	901	GTCCGTTGTAACCTCTTTACTGGCTTT	927
DB	901	GTCCGTTGTAACCTCTTTACTGGCTTT	927

RESULT 2

```

US/10/264-237-983/c
; Sequence 983, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Bitsee et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131PI
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn ver. 3.1
; SEQ ID NO 983
; LENGTH: 1732
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1568)..(1568)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1631)..(1631)

```

Db 473 AACGGCGCTTTTAGCGCGTTTTTATTTTCAACCTTATTTCCAGATACGTAACATCATC 414
Qy 901 GTCGGTGTAACTCTTTACTGGCTTT 927
Db 413 GTCCGKTGTAACTCTTTACTGGCTTT 387

RESULT 3
US-10-613-990A-19
; Sequence 19, Application US/10613990A
; Publication No. US20040180404A1
; GENERAL INFORMATION:
; APPLICANT: Ajinomoto Co., Inc.
; TITLE OF INVENTION: Method for Producing Target Substance by Fermentation
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/613,990A
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: JP 2002-203764
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Pantoea ananatis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (41)..(757)
US-10-613-990A-19

Query Match 54.9%; Score 508.6; DB 8; Length 759;
Best Local Similarity 80.6%; Pred. No. 1.2e-159;
Matches 607; Conservative 0; Mismatches 144; Indels 2; Gaps 1;

Qy 69 TTTGATTTAGTTGGC--AATTTAGTAGCAAAATGACAGACCCCGCATCTTATCGT 126
Db 7 TTTCAATTTAGTTGGCAAAATTTAGTAGCTAAACATGACAGACCCCGCATCTTATCGT 66

Qy 127 TGAACAGAGTTGGTAAACAGCAACACGTTGAAAAGTATTTTCGAAGCGGAAGGTATGA 186
Db 67 TGAACAGAGTTGGTACGCGGATACCTTAAAGCAATTTTGAAGCGGAAGGTATGT 126

Qy 187 TGTTTTCGAAGCGACAGATGGCGCGGAATGCAATCAGATCTCTCTGAATATGACATCAA 246
Db 127 CGTGTACGAAGCGACCGATGGTGACAGATGACACGAGTGTGTGACGCAATGATGTCAA 186

Qy 247 CCTGTGATCATGATATCAATCTCCGCGGTAAAGACGCTCTTCTGTTAGCGCGTGAAT 306
Db 187 TCTGTTTATATGACATCAATCTCGCGGTAAAGACGCGCTGTTACTGGCACGTTGAAT 246

Qy 307 GCGGAGCAGCGGAATGTTGGTTCATGTTCTGACTGCGCGTGACACGAAAGTCGATAA 366
Db 247 GCGTGAGCAAGCAATGTCGCAATGATGTTCTGACCGGACCGGATACGAAAGTCGATAA 306

Qy 367 AATTTCTCGGCTCGAATCGGTGACAGTACATACCAACCGTTTCAACCCCGTGA 426
Db 307 AATTTCTGGGCTGGAAATTTGGTGACAGACTACATTTACTAAGCGGTTTAAACCCAGCGA 366

Qy 427 ACTGACGATCTGTGACGCAACCTACTGTCGGTACCAATGATCTGGGTACTGTGACGGA 486
Db 367 ATTAACTATTCTGTGACGCTAACCTGCTGTTGGCACCACATGAAATTTGCGCTTACCCAAATGA 426

Qy 487 AGAAGCTCTAGCTGTGAAGCTACAAAGTTCAATGTTGGGAATCTGGGACATCAACAGCGG 546
Db 427 AGAGCGTCCGCAAGTTGAAGCTATAAGTTTCAACCGCTGGGAGCTGGACATCAACAGCGG 486

Qy 547 TTCTGTTGTCGCGCTGTATGGCGAGCAGTACAAAGCTGCGCGCAGCGAGTTCCCGGCCAT 606
Db 487 CTCACCTCATCAATCCCAAGGTGAGCAGTACAAAATGCGCGCAGTGAAGTTCCCGTCCAT 546

Qy 607 GCTTCACTTCTGTGAAAACCCAGGCAAAATTCAGTCCCGTGTGTAATCTGCTGAAGAAAT 666
Db 547 GCTGCACTTCTGCGAAAATCCCGGCAAGATTCAGACGCGGTGCTGATTTGCTGAAGAAAT 606

Qy 667 GACCGCGGTGAGCTGAAACCGCACGCGTACTGTAGACGTATCCCGCGTATTCG 726
Db 607 GACCGGAGCGGATCTCAAGCCACAGACCGTACTGTTGACGTGCAATCCGTCGTATCG 666
Qy 727 TAAACATTTTGAATCTACGCGGATACGCGGAATCATCGCCACCATTCACGGTGAAGG 786
Db 667 TAAACATTTTGAATCTCAAGCCAGATACCCCTGAAATCATCGCCACCATTCACGGGGAAGG 726
Qy 787 TTATCGCTTCTGCGGTGATCTGGAAGATTAATC 819
Db 727 TTATCGTTTCTGTGTGACCTGCAGGATTAAGC 759

RESULT 4

US-10-450-763-28743
; Sequence 28743, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 28743
; LENGTH: 1048
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (569)..(1021)
; OTHER INFORMATION: 98% homologous to Escherichia coli dye, accession number
; OTHER INFORMATION: M10044, Smith-Waterman Score=775.
US-10-450-763-28743

Query Match 48.3%; Score 447.4; DB 9; Length 1048;
Best Local Similarity 99.8%; Pred. No. 5.5e-139;
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 312 AGCAGCGGAATGTTGCGTTGATGTTCTCTGACTGGCCGTGACACGAAAGTCGATAAAATTC 371
Db 573 AACAGCGGAATGTTGCGTTGATGTTCTCTGACTGGCCGTGACACGAAAGTCGATAAAATTC 632

Qy 372 TCGGCTCGAATCGGTGCAGATGACTACATCACCACCGTTTCAACCCCGGTGAACCTGA 431
Db 633 TCGGCTCGAATCGGTGCAGATGACTACATCACCACCGTTTCAACCCCGGTGAACCTGA 692

Qy 432 CGATTCTGTGACGCAACCTACTGTCCCGTACATGAATCTGGGTACTGTGACGGAAGAAC 491
Db 693 CGATTCTGTGACGCAACCTACTGTCCCGTACATGAATCTGGGTACTGTGACGGAAGAAC 752

Qy 492 GTGCTAGCGTTGAAAGCTACAAAGTTCAATGTTGGGAATCTGGAATCAACAGCCGTTGCT 551
Db 753 GTGCTAGCGTTGAAAGCTACAAAGTTCAATGTTGGGAATCTGGAATCAACAGCCGTTGCT 812

Qy 552 TGATCGGCCCTGTGCGGCGAGTACAGCTGCGCGCAGCGAGTTCCGCGCATGCTTC 611
Db 813 TGATCGGCCCTGTGCGGCGAGTACAGCTGCGCGCAGCGAGTTCCGCGCATGCTTC 872

Qy 612 ACTTCTGTGAAAACCCAGGCAAAATTCAGTCCCGTGTGAACTGCTGAAGAAAATGACCG 671
Db 873 ACTTCTGTGAAAACCCAGGCAAAATTCAGTCCCGTGTGAACTGCTGAAGAAAATGACCG 932

Qy 672 GCGGTGAGCTGAAACCCGACGACCGTACTGTAGACGTGACGATCCGCGGTATTCGTAAC 731

Db 933 GCCGTGAGCTGAACCGCACGACCGTACTGTAGAGTGCAGGATCGCGGTATTCTGTAAC 992

QY 732 ATTTCGAATCTACGCGCGGANTACGCGGAA 760

Db 993 ATTTCGAATCTACGCGCGGANTACGCGGAA 1021

RESULT 5

US-10-795-159-693/c

; Sequence 683, Application US/10795159

; Publication No. US20050221439A1

; GENERAL INFORMATION:

; APPLICANT: BAKALETZ et al.

; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE

; FILE REFERENCE: 28335/38815A

; CURRENT APPLICATION NUMBER: US/10/795,159

; CURRENT FILING DATE: 2004-03-05

; PRIOR APPLICATION NUMBER: US 60/453,134

; PRIOR FILING DATE: 2003-03-06

; NUMBER OF SEQ ID NOS: 771

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 683

; LENGTH: 191996

; TYPE: DNA

; ORGANISM: H. influenzae

US-10-795-159-693

Query Match 34.6%; Score 320.6; DB 9; Length 191996;

Best Local Similarity 66.0%; Pred. No. 4.9e-95;

Matches 480; Conservative 0; Mismatches 244; Indels 3; Gaps 1;

QY 85 AATTAGGTAGCAACATGCAGACCCGCGACATCTTATCGTTGAAGACGAGTTCGTAAC 144

Db 55869 AGTATAAGGATGGAATAATGACTCTCCAAAATTCCTGTTGTAAGATGAATTTGTCC 55810

QY 145 ACGCAACACGTTGAAAGATATTTTGAAGCGGAAGCTATGATGTTTTCGAAGCGACAGA 204

Db 55809 TCGAATACGCTTAAGGATTTTGAAGCGGAGGATGATGTTTGAAGCGAGAAA 55750

QY 205 TGGCGGGAATATGATATGATATCTCTCTGGAATATGACATCAACCTGGTGATCATGATAT 264

Db 55749 TGGTGTGAAATGCAATCATATATATGCAATATCAATAATTAATTTGTTGTGATGATAT 55690

QY 265 CAATCTGCGGTGAGAACGCTTCTGTTAGCGGTGAACCTGCGGACGAGCGCAATGT 324

Db 55689 TAATTTACAGGCAAAACCGGCTTATATTGCAAGAGAACTCCGTTGAAGAAATTAAGCTT 55630

QY 325 TCGGTTGATGTTCCCTGACTGGCGGTGACAAACGAAGTCGATAAAATTCGGGCTCGAAAT 384

Db 55629 ACCTCTTATTTTAACTGGTTCGAGATATGAAGTGGTATAAAATTTAGGTCGGAAT 55570

QY 385 CGGTGAGATGATCATACACCAACCGGTTCAACCGCGGTGAACCTGACGATTCGTGACG 444

Db 55569 TGGTCCGACGATTAATTTAACCAACCTTTTAACCTAGAGAACTGACTATCCGCGCACG 55510

QY 445 CAACCTACTGTCGCTGACCATGAATCTGGTACTGTCTGCGGAAGAACGTCGTAGGTTGA 504

Db 55509 TAATCTATTGATCGCACAAATG---TCGCATCAAGAAAAGAAATATACATTTGGTCGAGA 55453

QY 505 AAGCTACAAGTTCAATGGTTGGGAATCGGACATCAACAGCCGTTTGGTTGATCGGCCCTGA 564

Db 55452 ATTCATCTGTTTAACTGGTTGGAAATTAGACCTAAATAGCCACAGTTTAATTTACACCAGA 55393

QY 565 TGGCGAGAGTACAGAGCTCCCGGCGACGAGTTCCGCGCCATGCTTCACTTCTGTGAANA 624

Db 55392 AGGACAAGAAATTCAACTTCTCGCAGTGAATTTCTGTCGAATGTTACATTTCTGTGAANA 55333

QY 625 CCAGGCAAAATTCAGTCCCGTCTGAACCTGCTGAAGAAATGACCGCGGTGAGCTGAA 684

Db 55332 TCAGGAAATTTGCAACCGCGGAGAAATTTCTGTAAGAAATATGACGGGACGAGAGTTAAA 55273

QY 685 ACCGACGACCGGTACTGTGTAGAGTGCAGATCCGCGGTATTTCGTAAACATTTTCCGAATCTAC 744

Db 55272 ACCTCAGGATCGTACCGTAGATGTCAAAATTCGACGTATCAGAAAACATTTTGAAGATCA 55213

QY 745 GCCGATACGCCGAAATCATCGCCACCATTCACGGTGAAGTTATCGCTTCTCGCGTGA 804

Db 55212 TCCAAATACTCCAAATATATCATTTAGCTATACATGGAGAGGCTATCGCTTTTGGGAGA 55153

QY 805 TCTGAA 811

Db 55152 TATTGAA 55146

RESULT 6

US-10-329-670-1/c

; Sequence 1, Application US/10329670

; Publication No. US20040018503A1

; GENERAL INFORMATION:

; APPLICANT: Fleischmann et al.

; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag

; TITLE OF INVENTION: Thereof, and Uses Thereof

; FILE REFERENCE: PB186P1

; CURRENT APPLICATION NUMBER: US/10/329,670

; CURRENT FILING DATE: 2002-12-24

; PRIOR APPLICATION NUMBER: US 09/643,990

; PRIOR FILING DATE: 2000-08-23

; PRIOR APPLICATION NUMBER: US 08/487,429

; PRIOR FILING DATE: 1995-06-07

; PRIOR APPLICATION NUMBER: US 08/426,787

; PRIOR FILING DATE: 1995-04-21

; NUMBER OF SEQ ID NOS: 1

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 1830121

; TYPE: DNA

; ORGANISM: Haemophilus influenzae

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (4747)..(4747)

; OTHER INFORMATION: n equals a, t, g or c

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (9921)..(9921)

; OTHER INFORMATION: n equals a, t, g or c

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (10150)..(10150)

; OTHER INFORMATION: n equals a, t, g or c

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (29298)..(29298)

; OTHER INFORMATION: n equals a, t, g or c

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (36543)..(36543)

; OTHER INFORMATION: n equals a, t, g or c

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (36551)..(36551)

; OTHER INFORMATION: n equals a, t, g or c

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (36636)..(36636)

; OTHER INFORMATION: n equals a, t, g or c

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (40808)..(40810)

; OTHER INFORMATION: n equals a, t, g or c

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (44416)..(44416)

; OTHER INFORMATION: n equals a, t, g or c

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (44905)..(44905)

```
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
```

```
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152530)..(152530)
```

```
Query Match          34.5%; Score 319.4; DB 7; Length 1830121;
Best Local Similarity 65.8%; Pred. No. 4.4e-94;
Matches 480; Conservative 0; Mismatches 246; Indels 3; Gaps 1;
QY      85 AATTTAGGTAGCAAAACATGCAGACCCCGCACATTCTTATCGTTGAAGACGAGTTGCTAAC 144
```

```
Db 936640 AGTATAGATGGAATGACTACTCCAAATAATCTCGTTGTTGAAGATGAATGTGCAC 936581
QY 145 ACGCAACACGTTGAAAGATATTTTCCAGCGGAAGCTATGATGTTTTCGAAGCGACAGA 204
Db 936580 TCGAAATACGCTTAAAGGATTTTTGAGCGGAAGATATGATGTTGTTGAAGCAGAAAA 936521
QY 205 TGGCGCGGAAATGCATCAGATCCTCTCTGMAATATGACATCAACCTGGTGATCATGGATAT 264
Db 936520 TGGTGTGAAATGCAATCATATATTGGCAATCATATAATTAATTTGGTTGTTGATGGATAT 936461
QY 265 CAATCTGCCGGTAAAGACGGTCTCTGTTAGCGGTGAATCTGCGCGAGCAGCGCAATGT 324
Db 936460 TAATTTACGAGCAAAACGGCTTATTATTGCAAGAGAACTCCCGTGAAGAAATTAAGCTT 936401
QY 325 TCGCTTGATGTTCTTGACTGCGCGTGACAAAGCTCGATAAAATTTCTCGGCTCGAAAT 384
Db 936400 ACCTCTTATTTTTTAACTGTCGAGATAATGAAGTGGATAAAATTTTAGGTTTGGAAAT 936341
QY 385 CGGTGCAGATGACTACATCACCACAAACCGTTCAACCGCGTGAACCTGACGATTCGTGCACG 444
Db 936340 TGGTCCGAGTACTATTTGACTAAGCCTTTTAAACCCTAGAGAACTGACTATCCCGGCACG 936281
QY 445 CAACCTACTGTCCCGTACCATGAATCTGGGTACTGTGACGGAAGAACGTGCTAGCGTTGA 504
Db 936280 TAATCTATTGCATCGTGCATG---CCGCATCAAGAAAAGAAATAACATTTGTCGAGA 936224
QY 505 AAGCTACAAGTTCAATGTTGGAACTGGACATCAACAGCCGTTTCGTTGATCGGCCCTGA 564
Db 936223 ATTCTATCGTTTTTAATGTTGGAAATTAGACCTAAATAGCCACAGTTTAAATTTACACCAGA 936164
QY 565 TGGCAGAGGTACAAGCTCCCGCAGCGAGTTCCGCGCCATGCTTCACTTCTGTGAAAA 624
Db 936163 AGGACAAGAATTCNAACCTTCTCCGAGTGAATTTCTGCAATGTTACATTTCTGTGAAAA 936104
QY 625 CCCAGCAAAATTCAGTCCCGTGTGNACTGCTGAAGAAAATGACCGCGCGTGAGCTGAA 684
Db 936103 TCCAGAAAATTTGCAACCGCGGAGAAATTTGCTGAAAAAATGACCGGACGAGTTAA 936044
QY 685 ACCGACGACCGTACTGTAGACGTGACGATCCGCCGTATTCGTAACATTTTCGAATCTAC 744
Db 936043 ACCTCAGGATCGTACGATGATGTACAAATTCGACGTATCAGAAAACATTTTGAAGATCA 935984
QY 745 GCGGATAGCCGGAATATCGCCACCATTCACGTTAGGTTATCGTTCGCGGTGA 804
Db 935983 TCCCAATACTCCAAATATCATATTGACTATACATGAGAGAGGCTATCGTTTGTGGAGA 935924
QY 805 TCTGGAAGA 813
Db 935923 TATTGATA 935915

RESULT 7
US-10-158-865-1/c
; Sequence 1. Application US/10158865
; Publication No. US20040203093A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
```

```
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
```

```
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a,t,c, or g

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152530)..(152530)

Query Match 34.5%; Score 319.4; DB 8; Length 180121;
Best Local Similarity 65.8%; Pred. No. 4.4e-94;
Matches 480; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

Qy 85 AATTTAGTAGCAACATGACGACCCGACATCTTATCTGTTGAGCGAGTGGTAAAC 144
Db 936640 AGTATAAGGATGGAAATGACTCTCCAAAAATTTCTCGTTGTAAGATGAATGTGAC 936581
Qy 145 ACGCAACACGTTGAAAAGTATTTTCGAAGCGAAGGCTATGATGTTTCGAAGCGACAGA 204
Db 936580 TCGAATACGCTTAAGGATTTTGAAGCGAAGGATGATGTTGTTGAAGCAGAAA 936521
Qy 205 TGGCGCGGAAATGCATCAGATCCTCTCTGAATATGACATCAACCTGGTGATCATGGATAT 264
Db 936520 TGGTGTGTAATGCATCATATATTGGCAAAATCATAATATTAAATTTGGTTGTGATGGATAT 936461
Qy 265 CAATCTGCGGGTAAGACGGTCTTCTGTTAGCGGTGAACTGCGGAGCGAGCGGAATGT 324
Db 936460 TAATTTACAGGCAAAAACGGCTTATTATTGGCAAGAGAACTCCGTGAAGAAATTAAGCTT 936401
Qy 325 TGGCTTGATGTTCTGACTGCGCGGTGCAACGAAGTGCATAAAATTTCTCGGCTCGAAAT 384
Db 936400 ACCTCTATTTTTTAAGTGTGCGAGATATGAATGATGATTAATTTTAGGTTGGAAT 936341
Qy 385 CGGTGCAGATGACTACATCACNAACCGTTCAACCCGCGTGAACCTGACGATTCGTGCACG 444
Db 936340 TGGTCCGATGACTATTGTAAGCCCTTTAAACCTAGAGAACTGACTATCCCGGCACG 936281
Qy 445 CAACCTACTGTCGCGTACCATGATCTGGGTACTGTCTAGCGAAGAACCTGCTAGGTTGA 504
Db 936280 TAATCTATTGTCATGCTGCAATG---CCGCATCAAGAAAAAGAAAAATACATTTGTCGAGA 936224
Qy 505 AAGCTTACAAGTTCAATGTTGGGAACCTGGACATCAACAGCCGTTCTGTTGATCGGCCCTGA 564
Db 936223 ATTCTATCGTTTTAATGTTGGAATTAGACTTAATAGCCACAGTTTAATTAACACAGA 936164
Qy 565 TGGCGAGCAGTACAAGCTGCGCGCGAGCGAGTTCCGCGCCATGCTTCACTTCTTGTGAAAA 624
```

```
Db 936163 AGGACAAGATTCAACTTCTCGCAGTGAATTTCTGTCATGTTACATTTCTGTGAAAA 936104
QY 625 CCCAGGCAAAATTCAGTCCGCTGCTGAACCTGCTGAAGAAATGACCGCGCTGAGCTGAA 684
Db 936103 TCCAGGAAATTTGCAACGCGCGAAGAATTGCTGAAAAAATGACCGGACGAGAGTTAAA 936044
QY 685 ACCGCACACCGTACTAGTAGAGTACGATCCGCGGTATTCGTAACATTTTCAATCTTAC 744
Db 936043 ACCTCAGGATCGTACCGTAGATGTACAAATTCGACGTATCAGAAACATTTTGAAGATCA 935984
QY 745 GCGCGATACGCGGAAATCATCGCCACCATTCACCGTGAAGCTTATCGTTCTTCCGGTGA 804
Db 935983 TCCCAATACTCCAAATATCAATATGACTATACATGATGAGAGGCTATCGTTTGTGGAGA 935924
QY 805 TCTGGAAGA 813
Db 935923 TATTGAATA 935915

RESULT 8
US-10-981-687-1/c
; Sequence 1, Application US/10981687
; Publication No. US20050131222A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; FILE REFERENCE: PB186P2C1D12
; CURRENT APPLICATION NUMBER: US/10/981,687
; PRIOR FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US 10/158,865
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40810)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4416)..(4416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
```


NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature

LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
Query Match 34.5%; Score 319.4; DB 9; Length 1830121;
Best Local Similarity 65.8%; Fred. No. 4.4e-94;
Matches 480; Conservative 0; Mismatches 246; Indels 3; Gaps 1;
QY 85 AATTTAGGTAGCAAAATGATGAGACCCCGACATTTCTTATCGTTGAAGACGAGTTCGTAAAC 144
DB 936640 AGTATAGGATGGAAATGACTACTCTCCAAAATTTCTCGTTGTTGAAGATGAATTTGTAC 936581
QY 145 ACGCAACACGTTGAAAAGTATTTTCGAAGCGGAAGGCTATGATGATTTTCGAAGCGACAGA 204
DB 936580 TCGAAATACGCTTAAAGGGATTTTGAAGCGGAAGGATATGATGTTGTTGAAGCGAATA 936521
QY 205 TGGCGCGGAATGATCAGATCCTCTCTGATATGACATCAACCTGATCATCATGATAT 264
DB 936520 TGGTGTGAAATGATCATATATTGGCAAAATCAATAATTAATTTGGTTGTGATGATAT 936461
QY 265 CAATCTCGCGGTAAAGACGGTCTTCTGTAGCGCTGAACTGCGCGGACGAGGCGCAATGT 324
DB 936460 TAATTTACGAGCAAAACGGCTTATTTATTTGCGAAGAACTCCGTTGAAGATTAAGCTT 936401
QY 325 TGGGTTGATGTTCTGACTGCGCGGTGCAACGAAAGTCGATATAAATTTCTGGGCTCGAAAT 384
DB 936400 ACCTCTTATTTTAACTGTCGAGATAATGAAGTGGATAAAATTTTAGGTTTGGAAAT 936341
QY 385 CGGTGAGATGACTACATCACAACCGTTTCAACCGCGGTGAACTGACGATTCGTGCGACG 444
DB 936340 TGGTGGCGATGACTATTTGACTAAGCCTTTTAAACCTTAGAAGACTGACTATCCGCGCAG 936281
QY 445 CAACCTACTGTCGGTACCATGAATCTGGGTACTGTCTAGCGAAGAACTCGTAGCGTTGA 504
DB 936280 TAATCTATTGATCGTGAATG---CCGCATCAAGAAAAGAAATAACATTTGGTCGAGA 936224
QY 505 AAGCTACAAAGTTCAATGTTGGGAATCTGGAATCAACAGCGGTTGTTGATGCGGCCCTGA 564
DB 936223 ATTCTATCGTTTAAATGTTGAAATTAGACTAAATAGCCACAGTTTAAATTTACACCGAGA 936164
QY 565 TGGCGAGCAGTACAAAGTGGCGCGGACGAGTTTCGGGCCATGCTTCACTTCTGTGAAA 624
DB 936163 AGGACAAAGATTCAAACTTCTCGCAGTGAATTTGTCGCAATGTTTACATTTCTGTGAAA 936104
QY 625 CCCAGGCAAAATTCAGTCCCGTGTGAACTGCTGAAGAAAATGACCGCGCGTGAAGCTGAA 684
DB 936103 TCCAGGAAAATTCGAAACGCGGAGAAATTCCTGAAAATAATGACCGGACGAGATTAA 936044
QY 685 ACCGACGACCGTACTGTAGACGTGACGATCCGCGGTATTCGTAACATTTGGAATCTAC 744
DB 936043 ACCTCAGGATCGTACCGTAGATGTCAAAATTCGACGTATCAGAAAACATTTTGAAGATCA 935984
QY 745 GCGGATACGCGGAAATCATCGCCACCATTCACGGTGAAGTTATCCGTTCTGCGGTGA 804
DB 935983 TCCCAATACTCCAAATATCATTTATGACTATATCATGAGAGAGGCTATCGTTTTGTGGAGA 935924
QY 805 TCTGGAAGA 813
DB 935923 TATTGATA 935915

RESULT 9
US-09-815-242-7046
Sequence 7046, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.

```
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7046
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(711)
US-09-815-242-7046

Query Match          34.1%; Score 316.2; DB 3; Length 711;
Best Local Similarity 66.2%; Pred. No. 6.3e-95;
Matches 472; Conservative 0; Mismatches 238; Indels 3; Gaps 1;

QY 101 ATGCAGACCCGCACATTCTTATCGTTGAAGCAGAGTTGGTAACACGCAACACGTTGAAA 160
Db 1 ATGACTACTCCAAAATTTCTCGTTGAGATGAATTTGTCACTCGAATACGCTTAAA 60

QY 161 AGTATTTTCGAAGCGGAAGGCTATGATGTTTTCGAAGCGACAGATGGCGCGGAATGCA 220
Db 61 GGGATTTTGAAGCGGAAGGATGATGTTTGAAGCAGAAAATGGTGTGAAATGCA 120

QY 221 CAGATCTCTCGAATATGACATCAACCTGGTGATCATGGATATCAATCTGCGGGTAAG 280
Db 121 CATATATTGGCAATCATTAATTAATTTGGTGTGATGGATTAATTTTACCAGGCAA 180

QY 281 AACGGTCTCTGTAGCGGTGAACCTCGCGAGCAGGCGAATGTTGCGTTGATGTTCTGT 340
Db 181 AACGGCTTATTTGGCAGAGAACTCCGTGAAGATTAAGCTTACCTCTTATTTTAA 240

QY 341 ACTGCGCGTGCAACGAAAGTCGATAAAATTTCTCGGCTCGAAATCGGTGCGAGATCACTAC 400
Db 241 ACTGGTCGAGATAATGAAGTGATAAAATTTTAGGTTTGGAATTTGGTGGCCGACTAT 300

QY 401 ATCACCACCGTTCAACCCGCGTGAACCTGACGATTCGTGACGCAACCTACTGTCGGT 460
Db 301 TTGACTAGCCCTTTAAACCCCTAGAGAACTGACTATCCGCGCACGTAATCTATTGCA 360

QY 461 ACCATGAATCTGGGTACTGTGACGCAAGAAACGTCGTAGCGTTGAAAGCTACAAGTTCAAT 520
Db 361 GCATG---CCGCATCAAGAAAAGAAATACATTTGGTCGAGAATCTATGTTTAA 417

QY 521 GGTGGGAATCGGACATCAACAGCCGTTGTTGATCGGCCCTCGAATCGGTGCGAGATCAAG 580
Db 418 GGTGGAAATTAGACCTAAATAGCCACAGTTTAAATACACAGAGGACAAAGAAATTCAAA 477

QY 581 CTGCGCGCAGGAGTTCCGCGCCATGCTTCACTTCTGTGAAAACCCAGGCAAAATTCAG 640
Db 478 CTTCTCGCAGTGAATTTTCGTGCAATGTTTACATTTCTGTGAAAATTCAGGAAAATTCGAA 537

QY 641 TCCCGTGCTGAACTGCTGAAGAAAATGACCGCGCTGAGCTGAAACCCGACCGACCGTACT 700
Db 641 GGGATTTTGAAGCGGAAGGATATGATGTTTGAAGCAGAAAATGGTGTGGAATGCA 120

Db 538 ACGCGCAAGAAATTGCTGAAAAAATGACGCGACGAGAGTTAAAACTCAGGATCGTACC 597
QY 701 GTAGACGTGAGATCCCGCGTATTCGTAACATTTGTAATCTACCGGATACGCGGAA 760
Db 598 GTAGATGTCAATTCGACGTTATCAGAAAACATTTTGAAGATCATCCCAATACTCCAAAT 657
QY 761 ATCATCGCCACCATTCACGGTGAAGGTTATCGCTTCTGCGGTGATCTGGAAGA 813
Db 658 ATCATTATGACTATACATGGAAGGCTATCGTTTTTGTGGAGATATTGAATA 710

RESULT 10
US-10-282-122A-22157
; Sequence 22157, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22157
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
US-10-282-122A-22157

Query Match          34.1%; Score 316.2; DB 7; Length 711;
Best Local Similarity 66.2%; Pred. No. 6.3e-95;
Matches 472; Conservative 0; Mismatches 238; Indels 3; Gaps 1;

QY 101 ATGCAGACCCGCACATTCTTATCGTTGAAGCAGAGTTGGTAACACGCAACGTTGAAA 160
Db 1 ATGACTACTCCAAAATTTCTCGTTTGAAGATGAATTTGTCACTCGAATACGCTTAAA 60

QY 161 AGTATTTTCGAAGCGGAAGGCTATGATGTTTTCGAAGCGACAGATGGCGCGGAATGCA 220
Db 61 GGGATTTTGAAGCGGAAGGATATGATGTTTGAAGCAGAAAATGGTGTGGAATGCA 120
```

Qy 221 CAGATCCTCTCTGAATATGACATCAACCTGGTATCATGGATATCAATCTCCCGGTAAG 280
Db |||||
Qy 121 CATATTTGGCAAAATATAATTAATTTGGTTGTGATGATTAATTAATTTACCGGCAAA 180
Db |||||
Qy 281 AACGGTCTCTCTGTTAGCGGTGAATCTGCGGAGCAGGCGAATGTTGCGTTGATGTTCTG 340
Db |||||
Qy 181 AACGGCTTATTATTGGCAAGAACTCCGTGAAGAAATTAAGCTTACCTCTTATTTTATTA 240
Db |||||
Qy 341 ACTGCCGTGACAAAGAACTGATATAAATTTCTCGGCTCGAAATCGGTGCGAGATGACTAC 400
Db |||||
Qy 241 ACTGGTCAGATAATGAAGTGATTAATTTAGTTTGGAAATTTGGTGGCGATGACTAT 300
Db |||||
Qy 401 ATCACAACCGTTCAACCGCGTGAACCTGAGATCTGTGACGCAACCTACTGTCCTG 460
Db |||||
Qy 301 TTGACTAAGCCCTTTTAAACCTAGAGAACTGACTATCCGCGCAGCTAATCTATTGCTAT 360
Db |||||
Qy 461 ACCATGAATCTGGGTACTGTGACGCAAGAACTGCTAGCTGTTGAAAGCTTACAAGTTCAAT 520
Db |||||
Qy 361 GCAATG---CCGATCAAGAAAGAAATATCATTTGGTCGAGAAATTTCTATCGTTTAAAT 417
Db |||||
Qy 521 GGTGGGAACCTGGAATCAACAGCCGTTCTGTTGATCGGCCCTGTATGGCGAGCAGTACAAG 580
Db |||||
Qy 418 GGTGGAAATAGACTTAATATAGCACAGTTTAATTAACACGAGGACAGAAATTCAAA 477
Db |||||
Qy 581 CTGCGCGGAGGAGTTCGCGGCCATGCTTCACTTCTGTGAAAACCCAGGCAAAATTCAG 640
Db |||||
Qy 478 CTTCCTCGAGTGAATTTCTGTCATGTTATCAATTTCTGTGAAAATCCAGGAAAATTCGAA 537
Db |||||
Qy 641 TCCCGTCTGACTCTGCAAGAAATGACCGCGCTGAGCTGAACCCGACGACCGTACT 700
Db |||||
Qy 538 AGCGCGAAGAAATGCTGAAAAAATGACCGGACAGAGTTAAAACTTCAGATCGTACC 597
Db |||||
Qy 701 GTAGACGTGACGATCCGCGCTATTCTGTAACATTTTCAATCTACCGCGGATACGCCGAA 760
Db |||||
Qy 598 GTAGATGTCACAAATTCGAGTATCAGAAACATTTTGAAGATCATCCCAATCTCCAAT 657
Db |||||
Qy 761 ATCATCGCAACATTCAGGTAAGGTTATCGCTTCTCGGTGATCTCGGAAGA 813
Db |||||
Qy 658 ATCATATGACTATATCATGGAGAGGCTATCGTTTTTGTGGAGATTTGAATA 710
Db |||||

RESULT 11
US-10-450-763-18226
; Sequence 18226, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIORITY FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 18226
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (183)..(698)
; OTHER INFORMATION: 92% homologous to Escherichia coli GABA PERMEASE (4-AMINO BUTYRATE TRANSPORT CARRIER) (GABA-AMINO BUTYRATE PERMEASE), accession number D90890, Smith-Waterman Score=810.
US-10-450-763-18226

Query Match 25.1%; Score 232.8; DB 9; Length 1110;
Best Local Similarity 98.8%; Pred. No. 9.4e-67;

Matches 245; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 544 CCGTTCGTTGATCGCCCTGATGCGAGCAGTA-CAAGCTCGCGCAGCAGATTCGCGG 602
Db |||||
Qy 761 CCGTTCGTTGATCGCCCTGATGCGGAGCAGTACCAAGCTCGCGCAGCAGATTCGCGG 820
Db |||||
Qy 603 CCATGCTTCACTTCTGTGAAAAACCCAGGCAAAATTCAGTCCCGTCTGAACTGTGAAAGA 662
Db |||||
Qy 821 CCATGCTTCACTTCTGTGAAAAACCCAGGCAAAATTCAGTCCCGTCTGAACTGTGAAAGA 880
Db |||||
Qy 663 AAATGACCGGCGGTGAGCTGAAACCGCAGCAGCTACTGTAGACGTGACGATCCGCGCTA 722
Db |||||
Qy 881 AAATGACCGGCGGTGAGCTGAAACCGCAGCAGCTACTGTAGACGTGACGATCCGCGCTA 940
Db |||||
Qy 723 TTGCTAAACATTTTGAATCTACGCGGATACGCGGAAATCATGCCCACCATTCACGCTG 782
Db |||||
Qy 941 TTGCTAAACATTTTGAATCTACGCGGATACGCGGAAATCATGCCCACCATTCACGCTG 1000
Db |||||
Qy 783 AAGGTTAT 790
Db |||||
Qy 1001 AAGTTTAT 1008
Db |||||

RESULT 12
US-10-795-159-291
; Sequence 291, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US/10/795,159
; PRIORITY FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 291
; LENGTH: 2125
; TYPE: DNA
; ORGANISM: H. influenzae
US-10-795-159-291

Query Match 17.4%; Score 161.6; DB 9; Length 2125;
Best Local Similarity 65.5%; Pred. No. 1.2e-42;
Matches 253; Conservative 0; Mismatches 129; Indels 4; Gaps 1;
Qy 426 AACTGACGATTCGTGCGAGCAACCTACTGTCCCGTACCATGAATCTGGGTACTGTGCGG 485
Db |||||
Qy 1 AACTGACTATTCGCGCAGCTAATCTATTGTCATCGCAC----AATGTCGATCAAGAAAAG 56
Db |||||
Qy 486 AAGAACGTCGTAGCGTTGAAAGCTACAAGTTCAATGTTGGAACTGGACATCAACAGCC 545
Db |||||
Qy 57 AAAATACATTTTGGTCGAGAAATTCATCGTTTAAATGTTGAAATTAGACCTAAATAGCC 116
Db |||||
Qy 546 GTTCGTTGATCGGCCCTGATGCGCAGCAGTAGTACAAGCTGCGCGCAGCAGATTCGCGCCA 605
Db |||||
Qy 117 ACAGTTTAAATTACACGAGGACAGAAATTCAAATCTCTCGCAGTGAATTCGTGCAA 176
Db |||||
Qy 606 TGCTTCACTTCTGTGAAAAACCCAGGCAAAATTCAGTCCCGTCTGAACTGTGAAAGAAA 665
Db |||||
Qy 177 TGTATACATTTCTGTGAAAAATCCAGGAAAATTCGAAACCGCGGAGAAATTCGTGAAAAA 236
Db |||||
Qy 666 TGACCGCGGTGAGCTGAAACCGCAGCAGCTACTGTAGTACGATGACGATCCGCGTATTC 725
Db |||||
Qy 237 TGACGGGACGAGAGTTAAACCTCAGGATCGTACCGTAGATGTCAATTCGACGTATCA 296
Db |||||
Qy 726 GTAAACATTTTGAATCTACGCGGATACGCGGAAATCATGCCACCATTCACGCTGAAG 785
Db |||||
Qy 297 GAAACATTTTGAAGATCATCAATCTCAATATCTCAATATCTATGCTATCATGAGAG 356
Db |||||
Qy 786 GTTATCGCTTCTCGGCTGATCTGAA 811
Db |||||

```
Db 357 GCTATCGCTTTTGGCGAGATATTGAA 382

RESULT 13
US-10-795-159-230/c
; Sequence 230, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US/10/795,159
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 230
; LENGTH: 2227
; TYPE: DNA
; ORGANISM: H. influenzae
US-10-795-159-230

Query Match 16.6%; Score 153.6; DB 9; Length 2227;
Best Local Similarity 66.8%; Pred. No. 6.3e-40;
Matches 219; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 85 AATTAGGTAGCAACATGACAGCCCGCACATCTTATCGTTGAAGACGAGTTGGTAAC 144
Db 353 AGTATAAGGATGGAATAATGACTACTCCAAAATTTCTCGTTGTTGAAGATGAAATTTGTCAC 294
QY 145 AGCGAACGTTGAAAGTATTTTCGAAGCGAAGCTATGATGTTTCGAAGCGACAGA 204
Db 293 TCGAATAACGTTAAAGGATTTTGAAGCGGAAGGATATGATGTTTGAAGCGAAGAAA 234
QY 205 TGGCGGGAATGCATCAGATCCTCTCTGAATATGACATCAACCTGGTGATCATGGATAT 264
Db 233 TGGTGTGAAATGCATCATATATTGGCAATCATATATTATTATTGTTGATGATAT 174
QY 265 CAATTCGCGGTGAAGACGGTCTTCTGTAGCGGTGAATCTGCGGACGAGCGCAATGT 324
Db 173 TAAATTACGAGCAAAAACGGCTTATTATTGGCAAGAGAACTCCCGTGAAGAAATTAAGCTT 114
QY 325 TCGTTGATGTTCTGACTGCGGCGTGACAAAGATCGATAAAATCTCGGCTCGAAT 384
Db 113 ACCTCTTATTTTAACTGGTCGAGATAATGAAGTGGATAAAATTTTAGGTCTGGAAT 54
QY 385 CGGTGCAGATGACTACATCACCACCAACCG 412
Db 53 TGGTCCGACGATTTATTNACCAACCG 26

RESULT 14
US-10-770-824-1
; Sequence 1, Application US/10770824
; Publication No. US20040198954A1
; GENERAL INFORMATION:
; APPLICANT: Campos, Manuel
; APPLICANT: Baarsch, Mary Jo
; APPLICANT: Rosey, Everett
; APPLICANT: Ankenbauer, Robert
; APPLICANT: Warren-Stewart, Lynn
; APPLICANT: Suiter, Brian
; APPLICANT: Keach, Robin
; TITLE OF INVENTION: NOVEL PROTEINS FROM ACTINOBACILLUS PLEUROPNEUMONIAE
; FILE REFERENCE: PC9854A
; CURRENT APPLICATION NUMBER: US/10/770,824
; CURRENT FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: US/09/418,980
; PRIOR FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
```

```
; LENGTH: 1018
; TYPE: DNA
; ORGANISM: Actinobacillus pleuropneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (272)..(787)
US-10-770-824-1

Query Match 11.6%; Score 107.4; DB 8; Length 1018;
Best Local Similarity 80.3%; Pred. No. 1.4e-24;
Matches 126; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 662 AAAATGACCGCGCGTGAAGCTGAAACCGCACGACCGTACTGTAGACGTGACGATCCGCGT 721
Db 68 AAAATGACCGCGTGTGAGCTAAACCGCATGACCGTACTGTGATGTGACGATTCGTCGT 127
QY 722 ATTGTAACAATTTGGAATCTTACCGCGGATACGCGGAAATCATCGCCACCATTCACCGT 781
Db 128 ATTGTAACAATTTGGAATCACCTTAATACCGGAAATCATTTGTAACCAATTCATGCT 187
QY 782 GAAGTTATCGCTTCTCGGCTGATCTGGAAGATTAAT 818
Db 188 GAAGTTATCGCTTCTCGGCTGATCTGGAAGATTAAT 224

RESULT 15
US-10-240-689-39/c
; Sequence 39, Application US/10240689
; Publication No. US2003017543A1
; GENERAL INFORMATION:
; APPLICANT: NG, Wee Chit
; TITLE OF INVENTION: Molecular Markers
; FILE REFERENCE: 6565-65001/RJP
; CURRENT APPLICATION NUMBER: US/10/240,689
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: SG 200002150-1
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39
; LENGTH: 11275
; TYPE: DNA
; ORGANISM: E. coli
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: AE 000201
US-10-240-689-39

Query Match 11.5%; Score 106.8; DB 6; Length 11275;
Best Local Similarity 51.3%; Pred. No. 8.8e-24;
Matches 327; Conservative 0; Mismatches 302; Indels 9; Gaps 3;

QY 102 TGCAGACCCCGCACATCTTATCGTTGAAGACGAGTTGGTAACACGCAACACGTTGAAA 161
Db 6729 TGATGCCACATCACATTTGTTTGAAGATGAGCCGTTACCCAGGCGGATTAACAAT 6670
QY 162 GTATTTTCGAAGCGGAAGCTATGATGTTTCGAAGCGACAGATGCGCGGAAATGCATC 221
Db 6669 CCTACTTCACTCAGGAGGGGTATACCGTTTCCGTTACGCGGAGCGGTCTGGGCTACGG 6610
QY 222 AGATCCTCTCTGAATATGACATCAACCTGGTGATCATGATATCAATCTGCCGGGTAA 281
Db 6609 AAATATCGAATCAATCGGTAGATTTAATTTCTGCTGGATATCACTTACCCGATGAA 6550
QY 282 ACGGCTTTCTGTAGCGGTGAATCTGCGGACGAGCGGAATGTTGCGTTGATGTTCTGA 341
Db 6549 ATGGCTGATGTTAAACCGCGCCCTGCGAGAACGCTCAACGCGTGGGATTTATTCGGTTA 6490
QY 342 CTGGCCGTTGACAAAGATTCGATATAAAATTCGCGCTCGAATCGTGCAGATGACTACA 401
Db 6489 CCGGACGCGGATCGGATTCGATGATGATGATGATGATGATGATGATGATGATGATG 6430
QY 402 TCACCAACCGTTCAACCCCGGTGAATTCGATGATGATGATGATGATGATGATGATGAT 461
```

[illegible]

Search completed: February 13, 2006, 15:57:23
Job time : 847 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2006, 02:34:44 ; Search time 203 Seconds
(without alignments)
515.134 Million cell updates/sec

Title: US-10-613-990A-32
Perfect score: 1230
Sequence: 1 MPTPHILIVEDLVTRNLK.....PEIIATIHGEGYRCGDLED 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_21.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1230	100.0	238	8	ADJ38403 Echerichia
2	1197	97.3	241	7	ABO62466 Klebsiell
3	1145	93.1	238	8	ADJ38391 Pantoea a
4	1138	92.5	239	6	ABM67466 Photorhab
5	1133	92.1	241	7	ADF05920 Bacterial
6	964.5	78.4	236	6	AAU35550 Haemophil
7	964.5	78.4	236	6	ABU30417 Protein e
8	778	63.3	176	4	ABG28752 Nucleo hum
9	458	37.2	278	7	ABO84242 Pseudomon
10	400	32.5	236	5	ABP26003 Streptoco
11	400	32.5	236	6	ABU46510 Protein e
12	400	32.5	721	7	ABO79547 Pseudomon
13	394	32.0	236	5	ABP26002 Streptoco
14	394	32.0	236	8	ADV88454 Streptoco
15	394	32.0	236	8	ADV81866 Streptoco
16	394	32.0	236	8	ADV79707 Streptoco
17	389.5	31.7	237	5	ABB48823 Listeria
18	389.5	31.7	237	6	ABU32830 Protein e
19	389.5	31.7	240	6	ABM70299 Photorhab
20	386.5	31.4	235	6	ABU18458 Protein e
21	386.5	31.4	235	6	ABU44248 Protein e
22	384.5	31.3	249	7	ADF06436 Bacterial
23	376	30.6	233	4	AAU33417 Enterococ
24	376	30.6	234	4	AAU35155 Enterococ

25	376	30.6	234	6	ABU14636	Abu14636 Protein e
26	376	30.6	238	6	ABU43704	Abu43704 Protein e
27	376	30.6	246	7	ADC94508	Adc94508 E. faeciu
28	376	30.6	261	6	ABU29650	Abu29650 Protein e
29	375	30.5	233	2	AAU06430	Aay06430 Staphyloc
30	375	30.5	233	3	AAU15430	Aab15430 S. aureus
31	375	30.5	233	4	AAU34180	Aau34180 Staphyloc
32	375	30.5	233	4	AAU00850	Aau00850 S. aureus
33	375	30.5	233	4	AAG81693	Aag81693 S. epider
34	375	30.5	233	6	ABU08234	Abu08234 Staphyloc
35	375	30.5	233	8	ADZ74033	Adz74033 Staphyloc
36	375	30.5	234	5	ABP38218	Abp38218 Staphyloc
37	375	30.5	234	8	ADS07613	Ads07613 Staphyloc
38	375	30.5	235	4	AAU37407	Aau37407 Staphyloc
39	375	30.5	235	4	AAU37081	Aau37081 Staphyloc
40	375	30.5	235	6	ABU16353	Abu16353 Protein e
41	375	30.5	235	6	ABM72823	Abm72823 Staphyloc
42	375	30.5	235	9	ADW94828	Adw94828 Prolifera
43	375	30.5	239	2	AAW89791	Aaw89791 Staphyloc
44	375	30.5	239	6	AAU42897	Abu42897 Protein e
45	374	30.4	294	7	ABO63249	AbO63249 Klebsiell

ALIGNMENTS

RESULT 1
ID ADJ38403 standard; protein; 238 AA.
XX
AC ADJ38403;
XX
DT 06-MAY-2004 (first entry)
XX
DE Echerichia coli ArcA (aerobic respiration control) protein SeqID32.
XX
KW gamma-proteobacterium; ArcA; aerobic respiration control;
KW ArcA protein function; chromosomal arcA gene; fermentation; L-amino acid;
KW L-lysine; L-glutamic acid.
XX
OS Escherichia coli.
XX
PN EP1382686-A1.
XX
PD 21-JAN-2004.
XX
PF 11-JUL-2003; 2003EP-00015911.
XX
PR 12-JUL-2002; 2002JP-00203764.
XX (AJIN) AJINOMOTO CO INC.
XX Ishikawa Y, Imaizumi A, Matsui K, Kojima H;
WPI; 2004-192975/19.
DR N-PSDB; ADJ38402.
XX
PT New bacterium useful for producing a target substance (especially L-lysine, L-glutamic acid or another L-amino acid) is modified so that the ArcA protein does not normally function.
XX
CC Claim 2; SEQ ID NO 32; 38pp; English.
XX
CC This invention relates to a novel gamma-proteobacterium able to produce a target substance such as an amino acid which is modified so that an ArcA (aerobic respiration control) protein does not normally function. The disruption to ArcA protein function is preferably due to disruption of a chromosomal arcA gene. The bacterium and method are useful in producing a target substance by fermentation. The target substance is preferably an L-amino acid, especially L-lysine or L-glutamic acid. Production of the target substance using the modified bacterium is more efficient. The present sequence is that of the E coli ArcA protein which was used in the method of the invention.

XX SQ Sequence 238 AA;
Query Match 100.0%; Score 1230; DB 8; Length 238;
Best Local Similarity 100.0%; Pred. No. 2.6e-127;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MOTPHILIVEDLVTRNTLKSIFEAEGYDVPEATDGAEMHQILSEYDINLVIMDINLPK 60
DB 1 MOTPHILIVEDLVTRNTLKSIFEAEGYDVPEATDGAEMHQILSEYDINLVIMDINLPK 60
QY 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIARNLISR 120
DB 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIARNLISR 120
QY 121 TNNLGTVSEERRSVESYKFNWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPGKIQ 180
DB 121 TNNLGTVSEERRSVESYKFNWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPGKIQ 180
QY 181 SRAELLKQMTGRELKHDPDRTVDVTTIRIRKHFESTPDTPEIIATIHGEGYRFGCDLED 238
DB 181 SRAELLKQMTGRELKHDPDRTVDVTTIRIRKHFESTPDTPEIIATIHGEGYRFGCDLED 238

RESULT 2
ABO62466
ID ABO62466 standard; protein; 241 AA.
XX AC ABO62466;
XX DT 29-JUL-2004 (first entry)
XX DE Klebsiella pneumoniae polypeptide seqid 8983.
XX KW Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX OS Klebsiella pneumoniae.
XX XX US6610836-B1.
XX PD 26-AUG-2003.
XX PF 27-JAN-2000; 2000US-00489039.
XX PR 29-JAN-1999; 99US-0117747P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Breton GL, Osborne M;
XX DR WPI; 2003-895346/82.
XX DR N-PSDB; ACH96017.
XX PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX preparing a vaccine composition against Klebsiella pneumoniae.
XX PS Disclosure; SEQ ID NO 8983; 932pp; English.
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
XX pneumoniae polypeptide. Also described are: a recombinant expression
XX vector comprising the nucleic acid, operably linked to a transcription
XX regulatory element; and a cell comprising the recombinant expression
XX vector. The nucleic acid is useful for preparing a vaccine composition
XX against Klebsiella pneumoniae. This is the amino acid sequence of a
XX Klebsiella pneumoniae polypeptide of the invention
XX SQ Sequence 241 AA;
Query Match 97.3%; Score 1197; DB 7; Length 241;
Best Local Similarity 97.1%; Pred. No. 1.2e-123;
Matches 231; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MOTPHILIVEDLVTRNTLKSIFEAEGYDVPEATDGAEMHQILSEYDINLVIMDINLPK 60
DB 4 MOTPHILIVEDLVTRNTLKSIFEAEGYDVPEATDGAEMHQILSEYDINLVIMDINLPK 63
QY 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIARNLISR 120
DB 64 NGLLLARELREQADVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIARNLISR 123
QY 121 TNNLGTVSEERRSVESYKFNWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPGKIQ 180
DB 124 TNNLGTVSEERRSVESYKFNWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPGKIQ 183
QY 181 SRAELLKQMTGRELKHDPDRTVDVTTIRIRKHFESTPDTPEIIATIHGEGYRFGCDLED 238
DB 184 SRAELLKQMTGRELKHDPDRTVDVTTIRIRKHFESTPDTPEIIATIHGEGYRFGCDLED 241
RESULT 3
ADJ38391
ID ADJ38391 standard; protein; 238 AA.
XX AC ADJ38391;
XX DT 06-MAY-2004 (first entry)
XX DE Pantoea ananatis Arca (aerobic respiration control) protein SeqID20.
XX KW gamma-proteobacterium; Arca; aerobic respiration control;
KW Arca protein function; chromosomal arca gene; fermentation; L-amino acid;
KW L-lysine; L-glutamic acid.
XX OS Pantoea ananatis.
XX XX EPI382686-A1.
XX PD 21-JAN-2004.
XX PF 11-JUL-2003; 2003EP-00015911.
XX PR 12-JUL-2002; 2002JP-00203764.
XX PA (AJIN) AJINOMOTO CO INC.
XX PI Ishikawa Y, Imaizumi A, Matsui K, Kojima H;
XX DR WPI; 2004-192975/19.
XX DR N-PSDB; ADJ38390.
XX PT New bacterium useful for producing a target substance (especially L-
XX lysine, L-glutamic acid or another L-amino acid) is modified so that the
XX Arca protein does not normally function.
XX PS Example 4; SEQ ID NO 20; 38pp; English.
XX This invention relates to a novel gamma-proteobacterium able to produce a
XX target substance such as an amino acid which is modified so that an Arca
XX (aerobic respiration control) protein does not normally function. The
XX disruption to Arca protein function is preferably due to disruption of a
XX chromosomal Arca gene. The bacterium and method are useful in producing a
XX target substance by fermentation. The target substance is preferably an L
XX -amino acid, especially L-lysine or L-glutamic acid. Production of the
XX target substance using the modified bacterium is more efficient. The
XX present sequence is that of the P ananatis Arca protein which was used in
XX the exemplification of the invention.
XX SQ Sequence 238 AA;
Query Match 93.1%; Score 1145; DB 8; Length 238;
Best Local Similarity 92.0%; Pred. No. 6.8e-118;
Matches 219; Conservative 11; Mismatches 8; Indels 0; Gaps 0;
QY 1 MOTPHILIVEDLVTRNTLKSIFEAEGYDVPEATDGAEMHQILSEYDINLVIMDINLPK 60
|||||

Db 1 MQTPHILIVEDLVTRNTLKSIFEAAGYVVVEATDGAEMHQVLTDNDVNLVIMDINLPGK 60
QY 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIARNLISR 120
Db 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIARNLISR 120
QY 121 TNLGTVSEERSVSYPKNGWELDINSRSLIGPDGEQYKLPSEFRAMLHFCENPGKIQ 180
Db 121 TNLPLPNEERRQVESYKFNWELDINSRSLINPNNGEQYKLPSEFRAMLHFCENPGKIQ 180
QY 181 SRALLKKMTGRELKHDPHRTVDVTIRIRKHFPSTPDTPETIATIHGEGYRFGCDLE 238
Db 181 TRADLLKKMTGRELKHDPHRTVDVTIRIRKHFPSTPDTPETIATIHGEGYRFGCDLQ 238

RESULT 4

ARM67466
ID ARM67466 standard; protein; 239 AA.

AC ARM67466;

XX 20-NOV-2003 (first entry)

DE Photorhabdus luminescens protein sequence #563.

XX
KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW detection; food; gene expression; plant; animal; microorganism; toxin;
KW antibiotic; biopesticide; virulence factor; disease model; plague;
KW whooping cough.

XX Photorhabdus luminescens.

OS WO200294867-A2.

PN 28-NOV-2002.

PD 07-FEB-2002; 2002WO-IB003040.

PF 07-FEB-2001; 2001FR-00001659.

PR (INSP) INST PASTEUR.

PA (CNRS) CNRS CENT NAT RECH SCI.

XX
PI Duchaud E, Tacurit S, Glaser P, Frangeul L, Kunst F, Danchin A;
PI Buchrieser C;

XX WPI; 2003-148459/14.

XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX Claim 2; SEQ ID NO 563; 1205pp; French.

XX The invention relates to the isolation of genes and their encoded
XX proteins from Photorhabdus luminescens. The isolated sequences are
XX sources of probes and primers for detecting the genome of P. luminescens
XX and related species; to study polymorphisms; for gene analysis and for
XX detection/amplification of the genes. Antibodies (Ab) raised against the
XX polypeptides encoded by the genes are used for detection/identification
XX of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
XX carry a gene-containing vector are used to select compounds that
XX modulate, regulate, induce or inhibit expression of the genes in plants,
XX animals or microorganisms other than P. luminescens and are able to alter
XX response or sensitivity to toxins and antibiotics produced by P.
XX recombinant production of the proteins, particularly toxins and
XX antibacterials useful as insecticides, bactericides and fungicides. The
XX genes, proteins, vectors containing the genes and Ab are also useful
XX therapeutically (to treat microbial infection by bacteria or fungi that
XX are sensitive to P. luminescens-encoded toxins or antibiotics) and as
XX biopesticides. Other uses of the genes and the proteins are as virulence
XX factors and for identifying targets of human diseases for which P.
XX luminescens is a model (particularly plague and whooping cough). This

CC sequence represents one of the isolated P. luminescens proteins

XX Sequence 239 AA;

Query Match 92.5%; Score 1138; DB 6; Length 239;

Best Local Similarity 93.2%; Pred. No. 4.1e-117;

Matches 221; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 MQTPHILIVEDLVTRNTLKSIFEAAGYVDVPEATDGAEMHQILSEYDINLVMIDINLPGK 60

Db 1 MQTPHILIVEDLVTRNTLKSIFEAAGYIVYEATDGSMMHILSNNDINLVMIDINLPGK 60

QY 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIARNLISR 120

Db 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIARNLISR 120

QY 121 TNLGTVSEERSVSYPKNGWELDINSRSLIGPDGEQYKLPSEFRAMLHFCENPGKIQ 180

Db 121 TNLNSVSEERRQVESYKFNWELDINSRSLISPAGEPYKLPSEFRAMLHFCENPGKIQ 180

QY 181 SRALLKKMTGRELKHDPHRTVDVTIRIRKHFPSTPDTPETIATIHGEGYRFGCDLE 237

Db 181 TRADLLKKMTGRELKHDPHRTVDVTIRIRKHFPSTPDTPETIATIHGEGYRFGCDLE 237

RESULT 5

ADP05920

ID ADP05920 standard; protein; 241 AA.

AC ADP05920;

XX 12-FEB-2004 (first entry)

XX Bacterial polypeptide #2033.

XX Proteus mirabilis infection; bacterial infection; antibacterial;
XX immunostimulant.

OS Proteus mirabilis.

XX US6605709-B1.

XX 12-AUG-2003.

XX 05-APR-2000; 2000US-00543681.

XX 09-APR-1999; 99US-0128706P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton GL;

XX WPI; 2003-895291/82.

XX N-PSDB; ADP01748.

XX New Proteus mirabilis polypeptides and polynucleotides, useful as
XX reagents for diagnosis of bacterial disease, as components of
XX antibacterial vaccines, as targets for antibacterial drugs, or as
XX biocontrol agents for plants.

XX Disclosure; SEQ ID NO 6205; 870pp; English.

XX The invention relates to new Proteus mirabilis polypeptides and
XX polynucleotides. The invention also relates to antibodies against the
XX polypeptides, methods for producing the polypeptides, a method of
XX generating vaccines for immunising an individual against P. mirabilis, a
XX method for evaluating a compound for the ability to bind a P. mirabilis
XX polypeptide and a method for screening test compounds for anti-bacterial
XX activity. The polypeptides and polynucleotides are useful as molecular
XX targets for diagnosing, preventing and treating pathological conditions
XX resulting from bacterial infection, as reagents for diagnosis of
XX bacterial diseases, as components of antibacterial vaccines, as targets
XX for antibacterial drugs or as bio-control agents for plants. This

CC sequence represents a *Proteus mirabilis* polypeptide of the invention.
XX
SQ Sequence 241 AA;
Query Match 92.1%; Score 1133; DB 7; Length 241;
Best Local Similarity 92.0%; Pred. No. 1.5e-116;
Matches 219; Conservative 8; Mismatches 11; Indels 0; Gaps 0;
QY 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVPEATDGAEMHQLSYDINLVMDINLPKG 60
DB 4 MQTPHILIVEDLVTRNTLKSIFEAEGYDVPEATDGAEMHQLSYDINLVMDINLPKG 63
QY 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIRARNLLSR 120
DB 64 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIRARNLLSR 123
QY 121 TNNLGTVSEERRSVSYKFNWELDINSRLIGPDGEQYKLPFRSEFRAMLHFCENPGKI 180
DB 124 TNNLANGTEERLVSYSKFNWELDINSRLISPTGEQYKLPFRSEFRAMLHFCENPGKI 183
QY 181 SRAELLKQMTGRELKHDPDRTVDVTTIRIRKHFESTPDTPETIATIHGEGYRFGDLE 238
DB 184 TRAEELLKQMTGRELKHDPDRTVDVTTIRIRKHFESTPDTPETIATIHGEGYRFGDLE 241
RESULT 6
AAU35550
ID AAU35550 standard; protein; 236 AA.
XX
AC AAU35550;
XX
DT 14-FEB-2002 (first entry)
XX
DE Haemophilus influenzae cellular proliferation protein #191.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX
OS Haemophilus influenzae.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR N-PSDB; AAS53409.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 11143; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also

CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 236 AA;
Query Match 78.4%; Score 964.5; DB 4; Length 236;
Best Local Similarity 76.2%; Pred. No. 6.9e-98;
Matches 183; Conservative 26; Mismatches 24; Indels 7; Gaps 2;
QY 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVPEATDGAEMHQLSYDINLVMDINLPKG 60
DB 1 MTTPKILVEDEIVTRNTLKGIFEAEGYDVPEAENGVMHILAHNINLVMDINLPKG 60
QY 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIRARNLLSR 120
DB 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIRARNLLSR 120
QY 121 TNNLGTVSEERRSV---ESYKFNWELDINSRLIGPDGEQYKLPFRSEFRAMLHFCENPG 177
DB 121 AM---PHQEKENTFGREFYRFGWKLDSLNSHSLITPEGQEFKLPFRSEFRAMLHFCENPG 176
QY 178 KIOSRAELLKQMTGRELKHDPDRTVDVTTIRIRKHFESTPDTPETIATIHGEGYRFGDLE 237
DB 177 KQTPREELLKQMTGRELKHDPDRTVDVTTIRIRKHFEDHPNTPNIMTIHNGEGYRFGDIE 236
RESULT 7
ABU30417
ID ABU30417 standard; protein; 236 AA.
XX
AC ABU30417;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #15944.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Haemophilus influenzae.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR N-PSDB; ACA34287.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 58341; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of

XX CC the 6213 antisense sequences given in the specification where expression

XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:

XX CC (1) a vector comprising a promoter operably linked to the nucleic acid

XX CC encoding a polypeptide whose expression is inhibited by the antisense

XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX CC polypeptide or its fragment whose expression is inhibited by the

XX CC antisense nucleic acid; (4) an antibody capable of specifically binding

XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

XX CC proliferation or the activity of a gene in an operon required for

XX CC the gene product or that has an activity against a biological pathway

XX CC required for proliferation, or that inhibits cellular proliferation; (8)

XX CC identifying a gene required for cellular proliferation or the biological

XX CC pathway in which a proliferation-required gene or its gene product lies

XX CC or a gene on which the test compound that inhibits proliferation of an

XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

XX CC compound's activity; (11) a culture comprising strains in which the gene

XX CC product is overexpressed or underexpressed; (12) determining the extent

XX CC to which each of the strains is present in a culture or collection of

XX CC strains; or (13) identifying the target of a compound that inhibits the

XX CC proliferation of an organism. The antisense nucleic acids are useful for

XX CC identifying proteins or screening for homologous nucleic acids required

XX CC for cellular proliferation to isolate candidate molecules for rational

XX CC drug discovery programs, or for screening homologous nucleic acids

XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

XX CC the target prokaryotic essential genes. Note: The sequence data for this

XX CC patent did not form part of the printed specification, but was obtained

XX CC in electronic format directly from WIPO at

XX CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 236 AA;

Query Match 78.4%; Score 964.5; DB 6; Length 236;

Best Local Similarity 76.2%; Pred. No. 6.9e-98;

Matches 183; Conservative 26; Mismatches 24; Indels 7; Gaps 2;

Qy 1 MCTPHILVDELVTRNTLKIFEAAGYDVFEATGAEHQILSEYDNLNVMIDNLPK 60

Db 1 MTPKILVVEIVTRNTLKIFEAAGYDVFEAENGVMHHILANHNINLVMDNLPK 60

Qy 61 NGLLARELRQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIRARNLLSR 120

Db 61 NGLLARELRBELSULPLTGRDNEVDKILGLEIGADDYITKPNPRELTIRARNLLHR 120

Qy 121 TWNLGTVSEERSV---ESYKFNWELDINSRSLIGPDGEQYKLPSEFRAMLHFCENPG 177

Db 121 AM----PHQEKENTGREFYRFGWKDLNLSLITPSGEQFKLPSEFRAMLHFCENPG 176

Qy 178 KIQSRAELLKQMTGRELKPHDRTVDVTTIRIRKHFESTPDTPEIATIHGEGYRPGDLR 237

Db 177 KLTREELLKQMTGRELKPHDRTVDVTTIRIRKHFEDHPTNPNIIMTIHGEGYRPGDIE 236

RESULT 8

ABG28752

ID ABG28752 standard; protein; 176 AA.

XX AC ABG28752;

XX XX 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #28743.

XX DE Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX XX

FN WO200175067-A2.

XX PD 11-OCT-2001.

XX XX 30-MAR-2001; 2001WO-US008631.

XX XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX XX (HYSE-) HVSEQ INC.

XX XX Dmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS92939.

XX XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX PS Claim 20; SEQ ID NO 59111; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 176 AA;

Query Match 63.3%; Score 778; DB 4; Length 176;

Best Local Similarity 93.8%; Pred. No. 2.1e-77;

Matches 152; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

Qy 70 REQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIRARNLLSRMTNLGTYS 129

Db 18 QKQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIRARNLLSRMTNLGTYS 77

Qy 130 ERRSVESYKFNWELDINSRSLIGPDGEQYKLPSEFRAMLHFCENPGKIQSRAELLKQM 189

Db 78 ERRSVESYKFNWELDINSRSLIGPDGEQYKLPSEFRAMLHFCENPGKIQSRAELLKQM 137

Qy 190 TGRELKHDRVTVDVTTIRIRKHFESTPDTPEIATIHGEGYR 231

Db 138 TGRELKHDRVTVDVTTIRIRKHFESTPDTPEIATIHGEGYR 175

RESULT 9

ABO84242

ID ABO84242 standard; protein; 278 AA.

XX AC ABO84242;

XX XX 29-JUL-2004 (first entry)

XX DE Pseudomonas aeruginosa polypeptide #16417.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
KW Pseudomonas aeruginosa.
OS US6551795-B1.
XX 22-APR-2003.
PD
XX 18-FEB-1999; 98US-00252991.
XX 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI WPI; 2003-615309/58.
XX N-PSDB; ABD17813.
DR
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.
PT
XX Disclosure; SEQ ID NO 32988; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a P. aeruginosa nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-P. aeruginosa drugs, as templates for recombinant production of P. aeruginosa-derived peptides or polypeptides, as target components for diagnosis and/or treatment of P. aeruginosa-caused infection, and in detection of P. aeruginosa sequences or other sequences of Pseudomonas species using biochip technology. Sequences ABO67826-ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html
XX
SQ Sequence 278 AA;

Query Match 37.2%; Score 458; DB 7; Length 278;
Best Local Similarity 39.4%; Pred. No. 1.3e-41;
Matches 95; Conservative 47; Mismatches 91; Indels 8; Gaps 3;
QY 4 PHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSE-----YDINLVINDINLP 58
DB 39 PRVLVVDVDPVIRELLQAYLGEGYDVLCAAGNAEQAEACLAECALHGQPVVELVLDIRLP 98
QY 59 GRNGLLRLAREQANVLMFTGRDNEVDKTLGLEIGADDYITKPFNPRLTIRARNLL 118
DB 99 GGGGTLTRELVRSEVGILLITGRNDEIDRIVLGECGADDYVTKPLNPRELVSAKNLI 158
QY 119 SRTMNL-GTVSEERSVSYKFNKGWELDINSRLIGPDGEOYKLPFRSEFRAMLHFCENPG 177
DB 159 RRVRAQASAGAPQALR--QFGDWLLDADRRLIDHAGNETLLTHGEFQLLGAFLRNSG 216
QY 178 KIQSAELLKQMTGRELPHDRTVDVYTRIRKHFESTPDPTPEITATTHGEGYRCGLDE 237
DB 217 HTLSRDQLMDQIRNREWLPSRSDIVLGRLRRLKRLDDPAEPQLITTHGAGYLFATAAS 276
QY 238 D 238
DB 277 D 277

RESULT 10
ABP26003

ID ABP26003 standard; protein; 236 AA.
XX
AC ABP26003;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 1182.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae; group A streptococcus; Streptococcus pyogenes; antibacterial; antiinflammatory; infection; vaccine; meningitis; gene therapy.
KW
XX Streptococcus pyogenes.
OS
XX WO200234771-A2.
XX
XX 02-MAY-2002.
PD
XX 29-OCT-2001; 2001WO-GB004789.
XX
XX 27-OCT-2000; 2000GB-00026333.
XX 24-NOV-2000; 2000GB-00028727.
XX 07-MAR-2001; 2001GB-00005640.
XX
XX (CHIR-) CHIRON SPA.
XX (GENO-) INST GENOMIC RES.
XX
XX Telford J, Masignani V, Margarit Y RosI, Grandi G, Fraser C;
PI Tettelin H;
XX
XX WPI; 2002-352536/38.
XX N-PSDB; ABN66634.
DR
XX New Streptococcus protein for the treatment or prevention of infection or disease caused by Streptococcus bacteria, such as meningitis, and for detecting a compound that binds to the protein.
PT
XX Claim 1; Page 3273; 4525pp; English.
PS
XX The invention relates to a protein (ABP25413-ABP30895) from group B streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in the specification. The proteins have antibacterial and antiinflammatory activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and antibodies that bind (I) are used in the manufacture of medicaments for the treatment or prevention of infection or disease caused by Streptococcus bacteria, particularly S. agalactiae and S. pyogenes. Nucleic acids encoding (I) are used to detect Streptococcus in a biological sample. (I) is used to determine whether a compound binds to (I). A composition comprising (I) or a nucleic acid encoding (I), may be used as a vaccine or diagnostic composition. The disease caused by Streptococcus that is prevented or treated may be meningitis. Nucleic acid encoding (I) may be used to recombinantly produce (I) and may be used in gene therapy. Antibodies to (I) are used for affinity chromatography, immunoassays, and distinguishing/identifying Streptococcus proteins
XX
SQ Sequence 236 AA;
Query Match 32.5%; Score 400; DB 5; Length 236;
Best Local Similarity 39.5%; Pred. No. 2.7e-35;
Matches 92; Conservative 45; Mismatches 82; Indels 14; Gaps 4;
QY 6 ILIIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVINDINLPKNGLL 65
DB 4 ILIVDDEKPIISDIKFNLTKEGYDVTAFDGREAVTIFEEBKPDLLIIDLMLPELDGLEV 63
QY 66 ARELREQANVLMFTGRDNEVDKTLGLEIGADDYITKPFNPRLTIRARNLLSRTMNLG 125
DB 64 ARKIRKSHVPIIMLSAKDSFVKIGLEIGADDYITKPFNRELLARVKAHLRTEYIE 123
QY 126 T-VSEERSVSYKFNKGWELDINSRLIGPD-----GEQYKLPFRSEFRAMLHFCENPG 177

CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences ABO67826-
CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX Sequence 721 AA;

Query Match 32.5%; Score 400; DB 7; Length 721;
Best Local Similarity 38.2%; Pred. No. 1.5e-34;
Matches 89; Conservative 44; Mismatches 92; Indels 8; Gaps 3;

QY 5 HILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKNGILL 64
Db HILIVDDDDREITELVGNVLLKKNGLRTTIVADGQRAFLKLEANTVDLIVLDMPPGDDGLL 540
QY 65 LARELREQANVA--LMFLTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARNLLSRTM 122
Db LCRELVRGKHKATPVLMLTARNDETDRIGLEMGADDYITKPFSAARELLARINAVLRTR 600
QY 123 NLG---TVSEERSRVESYKFNKGWELDINSRSLIGPDGEQYKLPKPRSEFRAMLHFCENPGKI 179
Db MLPNNLTVSESRLLI--GFGQWQLDTSARHLDDAGTIVVALSGAEYLLRVFLDHPQRV 657
QY 180 QSRALLKQMTGRELKHDPDRTVDVTRIRKHFESTPDTPEIATIHGEGYRF 232
Db LSRDQLNLQTGREADIFDRSIDLLVRLRQLRGDDAREPEYIKTVRSEGYVF 710

RESULT 13

ABP26002
ID ABP26002 standard; protein; 236 AA.

XX AC ABP26002;

XX 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 1180.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antinflammatory; infection; vaccine; meningitis; gene therapy.

XX OS Streptococcus agalactiae.

XX PN WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB004789.

XX PR 27-OCT-2000; 2000GB-00026333.

XX PR 24-NOV-2000; 2000GB-00028727.

XX PR 07-MAR-2001; 2001GB-00005640.

XX (CHIR-) CHIRON SPA.

XX PA (GENO-) INST GENOMIC RES.

XX Telford J, Maignani V, Margarit Y Rosi, Grandi G, Fraser C;

XX Tettelin H;

XX WPI; 2002-352536/38.

XX N-PSDB; ABN66633.

XX New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.

XX Claim 1; Page 3273; 4525pp; English.

XX

CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX Sequence 236 AA;

Query Match 32.0%; Score 394; DB 5; Length 236;
Best Local Similarity 39.6%; Pred. No. 1.2e-34;
Matches 90; Conservative 44; Mismatches 91; Indels 2; Gaps 2;

QY 6 ILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKNGILL 65
Db ILIVDDDEKPSIDIIKFNLTKEGYETATAFDGRAALVQVAEPQDLIILDLMLPELDGLEV 63
QY 66 ARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARNLLSRTMNLG 125
Db AKEVRKTSHPILMILSAKDSFVKVIGLEIGADDYITKPFNSRELLARVKALHRTENTIE 123
QY 126 T-VSEERSRVESYKFNKGWELDINSRSLIGPD-GEQYKLPKPRSEFRAMLHFCENPGKIQSRA 183
Db TAVAESQAQNASDITIGELQILPDATKRGKBEIEIETHREFELHHLATHIGQVMTRE 183
QY 184 ELLKQMTGRELKHDPDRTVDVTRIRKHFESTPDTPEIATIHGEGY 230
Db HLETVNGDYFDGVRVTRVRLREKIEDTPGRPEYILTRRGVGY 230

RESULT 14

ADV88454

ID ADV88454 standard; protein; 236 AA.

XX AC ADV88454;

XX 24-FEB-2005 (first entry)

XX Streptococcus agalactiae protein sequence, SEQ ID 848.

XX Antibacterial; Vaccine; bacterial infection.

XX OS Streptococcus agalactiae.

XX PN FR2824074-A1.

XX PD 31-OCT-2002.

XX PF 26-APR-2001; 2001FR-00005642.

XX PR 26-APR-2001; 2001FR-00005642.

XX (INSP) INST PASTEUR.

XX PA (CNRS) CNRS CENT NAT RECH SCI.

XX Glaser P, Rusniok C, Chevallier P, Frangeul L, Lalioui L;

XX PI Zouine M, Couve E, Buchrieser C, Poyart C, Trieu CP, Kunst F;

XX DR WPI; 2004-101891/11.

XX Genomic nucleotide sequences encoding polypeptides of Streptococcus
PT agalactiae for the development of vaccines, diagnostic tools, DNA chips

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2006, 02:38:39 ; Search time 44 Seconds
(without alignments)
520.445 Million cell updates/sec

Title: US-10-613-990A-32
Perfect score: 1230
Sequence: 1 MQTPHILIVEDLVTRNLTK.....PEIIATHGEGYFCGDLED 238

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1230	100.0	238	1 JYECR	dye resistance pro
2	1230	100.0	238	2 G91298	dye resistance pro
3	1230	100.0	238	2 A86140	dye resistance pro
4	1227	99.8	238	2 A1076	global response re
5	1145	93.1	238	2 AG0056	aerobic respiratio
6	1098	89.3	238	2 H82084	aerobic respiratio
7	964.5	78.4	236	2 I64099	aerobic respiratio
8	535	43.5	234	2 G82164	DNA-binding respon
9	479	38.9	230	2 A64841	torCAD operon tran
10	476	38.7	230	2 B85635	response transcrip
11	476	38.7	230	2 F90772	response transcrip
12	458	37.2	244	2 H83022	probable two-compo
13	427	34.7	238	2 A33608	two component resp
14	420	34.1	238	1 S28674	hypothetical prote
15	400	32.5	246	2 F83133	probable two-compo
16	398	32.4	245	2 D95868	probable response
17	392	31.9	238	2 T31680	back protein - Bac
18	389.5	31.7	237	2 A11110	two-component resp
19	384.5	31.3	237	2 A1472	two-component resp
20	384	31.2	247	2 D72228	response regulator
21	379	30.8	236	2 C84153	two-component resp
22	378.5	30.8	232	2 A83794	two-component resp
23	378	30.7	239	2 A80017	transcription regu
24	378	30.7	254	2 B87396	DNA-binding respon
25	375	30.5	235	2 B89760	response regulator
26	374	30.4	235	2 S65967	two-component resp
27	373.5	30.4	239	1 R8ECOR	osmosensor respons
28	373.5	30.4	239	2 A00998	two-component resp
29	373.5	30.4	239	2 G91159	response regulator

30	373.5	30.4	239	2 F86005	hypothetical prote
31	373.5	30.4	239	2 S01366	regulatory protein
32	373.5	30.4	253	2 AH3249	two component resp
33	373.5	30.4	253	2 S03760	virG protein - Agr
34	373	30.3	253	2 AG2368	two-component resp
35	372.5	30.3	239	2 S41887	ompR protein - Sal
36	371.5	30.2	245	2 AH0481	probable response
37	370	30.1	242	2 B83246	two-component resp
38	369.5	30.0	239	2 S70205	osmosensor respons
39	368	29.9	239	2 T06976	regulatory protein
40	367.5	29.9	241	2 JS0239	virG protein - Agr
41	365.5	29.7	267	2 A25519	virG protein - Agr
42	365	29.7	234	2 D95142	DNA-binding respon
43	365	29.7	251	2 B98010	response regulator
44	363.5	29.6	224	2 B97200	response regulator
45	358.5	29.1	240	2 C82042	transcription regu

ALIGNMENTS

RESULT 1

JYECR

dye resistance protein - Escherichia coli (strain K-12)

N/Alternate names: aerobic respiration control protein ArcA; protein dye negative regula

C/Species: Escherichia coli

C/Date: 28-May-1986 #sequence revision 28-May-1986 #text_change 09-Jul-2004

C/Accession: A03561; S56625; H65255

R/Drury, L.S.; Buxton, R.S.

J. Biol. Chem. 260, 4236-4242, 1985

A/Title: DNA sequence analysis of the dye gene of Escherichia coli reveals amino acid ho

A/Reference number: A92515; MUID:85157583; PMID:2984198

A/Accession: A03561

A/Molecule type: DNA

A/Residues: 1-238 <DR>

A/Cross-references: UNIPROT:P03026; UNIPARC:UPI0000125D8C; GB:M10044; NID:g145817; PIDN:

R/Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.

Nucleic Acids Res. 23, 2105-2119, 1995

A/Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.

A/Reference number: S56314; MUID:95334362; PMID:7610040

A/Accession: S56625

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-238 <BUR>

A/Cross-references: UNIPARC:UPI0000125D8C; EMBL:U14003; NID:g1263172; PIDN:AAA97297.1; F

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994

R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A/Title: The complete genome sequence of Escherichia coli K-12.

A/Reference number: A64720; MUID:97426617; PMID:9278503

A/Accession: H65255

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-238 <BLAT>

A/Cross-references: UNIPARC:UPI0000125D8C; GB:AB000510; GB:U00096; UMGP:b4401; NID:g1790

A/Experimental source: strain K-12, substrain MG1655

C/Genetics:

A/Gene: arcA; dye; fexA; msp; seg; sfrA

A/Map position: 100 min

C/Function:

A/Description: regulatory protein, related to ompR, controlling resistance to dyes; requ

A/Note: may also be involved in the osmoregulation of envelope proteins

C/Superfamily: ompR protein; response regulator homology

C/Keywords: DNA binding; osmoregulation; phosphoprotein; transcription regulation

F:6-114/Domain: response regulator homology <RRH>

F:54/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 100.0% Score 1230; DB 1; Length 238;

Best Local Similarity 100.0%; Pred. No. 1.1e-85;

Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQTPHILIVEDLVTRNLTKSIFEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPCK 60

```
|||||
Db 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
QY 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDDYTKPFPNPRELTIARNLISR 120
Db 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDDYTKPFPNPRELTIARNLISR 120
QY 121 TNNLGTVSEERSVESYKFNKGWELDINSRLIGPDGEQYKLPFRSEFRAMLHFCENPGKIQ 180
Db 121 TNNLGTVSEERSVESYKFNKGWELDINSRLIGPDGEQYKLPFRSEFRAMLHFCENPGKIQ 180
QY 181 SRAELLKMTGRELKPHDRTVDVTIRIRKHFESTPDTPEIIATIHGEGYRFGCDLED 238
Db 181 SRAELLKMTGRELKPHDRTVDVTIRIRKHFESTPDTPEIIATIHGEGYRFGCDLED 238

RESULT 2
G91298
dye resistance protein arcA [similarity] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: G91298
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-238 <HAV>
A:Cross-references: UNIPROT:P03026; UNIPARC:UPI0000125D8C; GB:BA0000007; PIDN:BAB38782.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: EC85359
C:Superfamily: ompR protein; response regulator homology

Query Match 100.0%; Score 1230; DB 2; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.1e-89;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
Db 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
QY 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDDYTKPFPNPRELTIARNLISR 120
Db 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDDYTKPFPNPRELTIARNLISR 120
QY 121 TNNLGTVSEERSVESYKFNKGWELDINSRLIGPDGEQYKLPFRSEFRAMLHFCENPGKIQ 180
Db 121 TNNLGTVSEERSVESYKFNKGWELDINSRLIGPDGEQYKLPFRSEFRAMLHFCENPGKIQ 180
QY 181 SRAELLKMTGRELKPHDRTVDVTIRIRKHFESTPDTPEIIATIHGEGYRFGCDLED 238
Db 181 SRAELLKMTGRELKPHDRTVDVTIRIRKHFESTPDTPEIIATIHGEGYRFGCDLED 238

RESULT 3
A86140
dye resistance protein arcA [similarity] - Escherichia coli (strain O157:H7, substrain B
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: A86140
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: A86140
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-238 <STO>
```

```
A:Cross-references: UNIPROT:P03026; UNIPARC:UPI0000125D8C; GB:AE005174; NID:g12519433; F
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: arcA
C:Superfamily: ompR protein; response regulator homology

Query Match 100.0%; Score 1230; DB 2; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.1e-89;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
Db 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
QY 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDDYTKPFPNPRELTIARNLISR 120
Db 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDDYTKPFPNPRELTIARNLISR 120
QY 121 TNNLGTVSEERSVESYKFNKGWELDINSRLIGPDGEQYKLPFRSEFRAMLHFCENPGKIQ 180
Db 121 TNNLGTVSEERSVESYKFNKGWELDINSRLIGPDGEQYKLPFRSEFRAMLHFCENPGKIQ 180
QY 181 SRAELLKMTGRELKPHDRTVDVTIRIRKHFESTPDTPEIIATIHGEGYRFGCDLED 238
Db 181 SRAELLKMTGRELKPHDRTVDVTIRIRKHFESTPDTPEIIATIHGEGYRFGCDLED 238

RESULT 4
AF1076
Global response regulator arcA [imported] - Salmonella enterica subsp. enterica serovar
C:Species: Salmonella enterica subsp. enterica serovar typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AF1076
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AF1076
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-238 <PAR>
A:Cross-references: UNIPARC:UPI000005A9F8; GB:AL513382; PIDN:CAD03429.1; PID:g16505695;
C:Genetics:
A:Gene: arcA
C:Superfamily: ompR protein; response regulator homology

Query Match 99.8%; Score 1227; DB 2; Length 238;
Best Local Similarity 99.6%; Pred. No. 1.8e-89;
Matches 237; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
Db 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
QY 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDDYTKPFPNPRELTIARNLISR 120
Db 61 NGLLLARELREQANVALMFLTGRDNEVDKILGLEIGADDDYTKPFPNPRELTIARNLISR 120
QY 121 TNNLGTVSEERSVESYKFNKGWELDINSRLIGPDGEQYKLPFRSEFRAMLHFCENPGKIQ 180
Db 121 TNNLGTVSEERSVESYKFNKGWELDINSRLIGPDGEQYKLPFRSEFRAMLHFCENPGKIQ 180
QY 181 SRAELLKMTGRELKPHDRTVDVTIRIRKHFESTPDTPEIIATIHGEGYRFGCDLED 238
Db 181 SRAELLKMTGRELKPHDRTVDVTIRIRKHFESTPDTPEIIATIHGEGYRFGCDLED 238

RESULT 5
AG0056
```

aerobic respiration control protein [imported] - Yersinia pestis (strain C092)
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
 C:Accession: AG0056
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Li, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AG0056
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-238 <KUR>
 A:Cross-references: UNIPROT:Q8ZIN7; UNIPARC:UPI00000DCC6B; GB:AL590842; PIDN:CAC89314.1;
 C:Genetics:
 A:Gene: arca
 C:Superfamily: ompR protein; response regulator homology

Query Match 93.1%; Score 1145; DB 2; Length 238;
 Best Local Similarity 92.9%; Pred. No. 5.3e-83;
 Matches 221; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MQTPHILVEDELVTRNTLKSIFEAEGYDVFEATDGAEMHQILSYDINLVIMDINLPGK 60
 Db 1 MQTPHILVEDELVTRNTLKSIFEAEGYDVFEATDGAEMHQILSYDINLVIMDINLPGK 60

Qy 61 NGLLLARELREQANVALMFLTRGRDNEVDKILGLEIGADDYITKPPNPRELTIRARNLLSR 120
 Db 61 NGLLLARELREQANVALMFLTRGRDNEVDKILGLEIGADDYITKPPNPRELTIRARNLLSR 120

Qy 121 TMNLGTVSEERSVSYKFNWELINDSRSLIGPDGEQYKLPKRSFRAMLHFCENPGKI 180
 Db 121 TMNLGTVSEERSVSYKFNWELINDSRSLIGPDGEQYKLPKRSFRAMLHFCENPGKI 180

Qy 121 TMNLSSVGEERLVSYSKFNWELINDSRSLIGPDGEQYKLPKRSFRAMLHFCENPGKI 180
 Qy 181 SRABELLKMTGRELKPHDRTVDVTRIRKHFESTPDTPPEIATIHGEGYRFGCDLED 238
 Db 181 SRGELLKMTGRELKPHDRTVDVTRIRKHFESTPDTPPEIATIHGEGYRFGCDLED 238

RESULT 6
 H82084
 aerobic respiration control protein FexA VC2368 [imported] - Vibrio cholerae (strain N16)
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C:Accession: H82084
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, J.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: AB2035; MUID:20406833; PMID:10952301
 A:Accession: H82084
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-238 <HEI>
 A:Cross-references: UNIPROT:Q9KJ9; UNIPARC:UPI00000C329F; GB:AE004307; GB:AE003852; NID
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC2368
 A:Map position: 1
 C:Superfamily: ompR protein; response regulator homology

Query Match 89.3%; Score 1098; DB 2; Length 238;
 Best Local Similarity 88.2%; Pred. No. 2.7e-79;
 Matches 210; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

Qy 1 MQTPHILVEDELVTRNTLKSIFEAEGYDVFEATDGAEMHQILSYDINLVIMDINLPGK 60
 Db 1 MQTQILLVEDEQVTRNTLKSIFEAEGYDVFEATDGAEMHQILSYDINLVIMDINLPGK 60

Qy 61 NGLLLARELREQANVALMFLTRGRDNEVDKILGLEIGADDYITKPPNPRELTIRARNLLSR 120

Db 61 NGLLLARELREQANVALMFLTRGRDNEVDKILGLEIGADDYITKPPNPRELTIRARNLLSR 120
 Qy 121 TMNLGTVSEERSVSYKFNWELINDSRSLIGPDGEQYKLPKRSFRAMLHFCENPGKI 180
 Db 121 SMHAGTQGEKRSVEKYVFNWELINDSRSLIGPDGEQYKLPKRSFRALLHFCENPGKI 180
 Qy 181 SRABELLKMTGRELKPHDRTVDVTRIRKHFESTPDTPPEIATIHGEGYRFGCDLED 238
 Db 181 TRADLLKMTGRELKPHDRTVDVTRIRKHFESVSGTPEIATIHGEGYRFGCDLED 238

RESULT 7
 I64099
 aerobic respiration control protein arca homolog - Haemophilus influenzae (strain Rd KW2)
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
 C:Accession: I64099
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: I64099
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-236 <TIGR>
 A:Cross-references: UNIPROT:P44918; UNIPARC:UPI0000125D8D; GB:U32770; GB:L42023; NID:915
 C:Superfamily: ompR protein; response regulator homology
 C:Keywords: DNA binding; phosphoprotein; transcription regulation
 F:6-114/Domain: response regulator homology <RRH>
 F:54/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 78.4%; Score 964.5; DB 2; Length 236;
 Best Local Similarity 76.2%; Pred. No. 8.8e-69;
 Matches 183; Conservative 26; Mismatches 24; Indels 7; Gaps 2;

Qy 1 MQTPHILVEDELVTRNTLKSIFEAEGYDVFEATDGAEMHQILSYDINLVIMDINLPGK 60
 Db 1 MTTPKILVEDELVTRNTLKSIFEAEGYDVFEATDGAEMHQILSYDINLVIMDINLPGK 60

Qy 61 NGLLLARELREQANVALMFLTRGRDNEVDKILGLEIGADDYITKPPNPRELTIRARNLLSR 120
 Db 61 NGLLLARELREELSPLIFLTRGRDNEVDKILGLEIGADDYITKPPNPRELTIRARNLLSR 120

Qy 121 TMNLGTVSEERSV---ESYKFNWELINDSRSLIGPDGEQYKLPKRSFRAMLHFCENPG 177
 Db 121 AM----PHQEKENTFGREFYRFGNWKLDNLSLITPEGQBFKLPKRSFRAMLHFCENPG 176

Qy 178 KIQSRAELLKMTGRELKPHDRTVDVTRIRKHFESTPDTPPEIATIHGEGYRFGCDLE 237
 Db 177 KIQTBRELLKMTGRELKPHDRTVDVTRIRKHFEDHPNTPNITMTHGEGYRFGCDIE 236

RESULT 8
 G82164
 DNA-binding response regulator TorR VC1719 [imported] - Vibrio cholerae (strain N16961)
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C:Accession: G82164
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, J.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: AB2035; MUID:20406833; PMID:10952301
 A:Accession: G82164
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-234 <HEI>
 A:Cross-references: UNIPROT:Q9KRC3; UNIPARC:UPI00000C30AD; GB:AE004250; GB:AE003852; NID
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: F90772

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-230 <HAV>

A;Cross-references: UNIPROT:P58357; UNIPARC:UPI0000137178; GB:BA000007; PIDN:BAB34573.1;

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: ECs1150

C;Superfamily: ompR protein; response regulator homology

Query Match 38.7%; Score 476; DB 2; Length 230;

Best Local Similarity 42.5%; Pred. No. 2.7e-30;

Matches 99; Conservative 51; Mismatches 75; Indels 8; Gaps 4;

Qy 5 HILIVEDLVTRNTLKSIFAEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPKNGLL 64

Db 4 HIVIVEDPVTQARLOSFTQGYTVSVTASGAGREIMQVQVDLLDINLPDENGIM 63

Qy 65 LARELREQANVALMFLTGRDNEVDKILGLETGADDDYITKPFNPRLTIRARNLLSRWNL 124

Db 64 LTRALRSTVGIIILVTGRSDRIDRVGLEMGADDDYTKPLELRELVVVRKNLLWR-IDL 122

Qy 125 GTVSEERSVSYKFNKGWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPGKIQSRAE 184

Db 123 ARQAQPYTQDNCYPAGYCLNV-SRHTLERDGEPIKLTAEYEMLVAFVTPNGEILSR 181

Qy 185 LLKQMTGRELK-PHRTVDVTIRIRKHFESTPDPEIATIHGEGYRFGDGL 236

Db 182 LLRLMSARVENPDURTVDLIRLRHKLISA-----DLVTQHGEGYFLAADV 229

RESULT 12

H83022

probable two-component response regulator PA4983 [imported] - Pseudomonas aeruginosa (st

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C;Accession: H83022

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: H83022

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-244 <STO>

A;Cross-references: UNIPROT:Q9HUI2; UNIPARC:UPI00000C5EAC; GB:AE004911; GB:AE004091; NID

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA4983

C;Superfamily: ompR protein; response regulator homology

Query Match 37.2%; Score 458; DB 2; Length 244;

Best Local Similarity 39.4%; Pred. No. 7.6e-29;

Matches 95; Conservative 47; Mismatches 91; Indels 8; Gaps 3;

Qy 4 PHILIVEDLVTRNTLKSIFAEAGYDVFEATDGAEMHQILSE-----YDINLVIMDINLP 58

Db 5 PRVLVDDPVTIRELLQAYLGEEDYVLCAGNAEQAEACALHAGQFVELVLDIRLP 64

Qy 59 GNGGLLARELREQANVALMFLTGRDNEVDKILGLETGADDDYITKPFNPRLTIRARNLL 118

Db 65 GKGGLTLRELVRSEVGIILITGRNDEIDRVGLECGADDDYVTKPLNPRLEVSRAKLI 124

Qy 119 SRTMNL-GTVSEERSVSYKFNKGWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPG 177

Db 125 RRVHAQASAGPARQALR-QFGDWLLDADRRLIDHAGNETLLTHGEFQLLAGPLRNSG 182

Qy 178 KIQSRAELLKQMTGRELPHDRTVDVTIRIRKHFESTPDPEIATIHGEGYRFGDGL 237

Db 183 HTLSRQDLMDQIRNREWLPSRSDIDLVRRLRRKLDDPAPBQLIITIHGAGYLTAAAS 242

Qy 238 D 238

Db 243 D 243

RESULT 13

AF3608

two component response regulator [imported] - Brucella melitensis (strain 16M)

C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C;Accession: AF3608

R;DelVecchio, V.G.; Kapratl, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens

A;Reference number: AD3252; PMID:11756688

A;Accession: AF3608

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-238 <KUR>

A;Cross-references: UNIPROT:Q8YBU5; UNIPARC:UPI0000058604; GB:AE008918;

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BME10791

A;Map position: II

C;Superfamily: ompR protein; response regulator homology

Query Match 34.7%; Score 427; DB 2; Length 238;

Best Local Similarity 40.3%; Pred. No. 2e-26;

Matches 94; Conservative 44; Mismatches 93; Indels 2; Gaps 1;

Qy 5 HILIVEDLVTRNTLKSIFAEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPKNGLL 64

Db 6 HILIVDDKDIRDLLHFEKRGHVSITACNGDEMLVLSRTPIDLVLDVMLPKKSIE 65

Qy 65 LARELREQANVALMFLTGRDNEVDKILGLETGADDDYITKPFNPRLTIRARNLLSRWNL 124

Db 66 ICQDVRRTSRVPIIMLTALADADKILGLETGADDDYIAKPPDPRELLARIVLRFE-- 123

Qy 125 GTVSEERSVSYKFNKGWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPGKIQSRAE 184

Db 124 GNRSPQALTOIYRFAGTCLDCARRLTSPHDVVELTAEFNLLEAFVKSQHLSRSDQ 183

Qy 185 LLKQMTGRELPHDRTVDVTIRIRKHFESTPDPEIATIHGEGYRFGDGL 237

Db 184 LMENAGQAVGYDRSDVLLISRLRKKLEDDPCAPKLILITIRGGYQFGHEVE 236

RESULT 14

S28674

hypothetical protein 2 - Rhizobium sp. (strain IC3342)

C;Species: Rhizobium sp.

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C;Accession: S28674

R;Upadhyaya, N.M.; Scott, K.F.; Tucker, W.T.; Watson, J.M.; Dart, P.J.

Mol. Plant Microbe Interact. 5, 129-143, 1992

A;Title: Isolation and characterization of Rhizobium (IC3342) genes that determine leaf

A;Reference number: S28673; MUID:92314428; PMID:1319772

A;Accession: S28674

A;Molecule type: DNA

A;Residues: 1-238 <UPA>

A;Cross-references: UNIPROT:Q53167; UNIPARC:UPI00000B7CA9; EMBL:M38698; NID:gl522257; PID

C;Superfamily: ompR protein; response regulator homology

C;Keywords: phosphoprotein

F:4-111/Domain: response regulator homology <RRH>

F:51/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 34.1%; Score 420; DB 1; Length 238;

Best Local Similarity 38.4%; Pred. No. 7.2e-26;

Matches 89; Conservative 49; Mismatches 88; Indels 6; Gaps 3;

Qy 4 PHILIVEDLVTRNTLKSIFAEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPKNGLL 63

```
Db      2 PHLIIVDDDDPRIRSMISRYLEDEGFRVLAENISQLRRLVSP-SVDLVLLDGLPGDGNGL 60
Qy      64 LLARELREQANVALMFLTGRNEVDKILGLIGADDYITKFPNPRELTIRARNLL---SR 120
Db      61 ELAREIDANFRVPTTIIVSGRODDVDRIIGLEMGADDYVSKPFLNRELLARVRSVLRSQR 120
Qy      121 TNLGLTGVSERRSVESYKFNWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPGKIQ 180
Db      121 AMPLGPAQKATAGI--FHPDGWMLDADRQLTSTAGQTIELTTGEFDLLMVVTFHGRVL 178
Qy      181 SRAELKKMTGRELKPHDRTVDVTIRIRKHFEPTDPTPEIATIHGEYRF 232
Db      179 TRDFLLDQTRGTREAFDRAIDVQVTRLRAKVDEDDGDPRIKSVRGAGYVF 230

RESULT 15
F83133
Probable two-component response regulator PA4101 [imported] - Pseudomonas aeruginosa (st
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F83133
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bz
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83133
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-246 <STO>
A:Cross-references: UNIPROT:Q9HWS7; UNIPARC:UPI00000C5COF; GB:AE004826; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4101
C:Superfamily: ompR protein; response regulator homology

Query Match      32.5%; Score 400; DB 2; Length 246;
Best Local Similarity 38.2%; Pred. No. 2.8e-24;
Matches 89; Conservative 44; Mismatches 92; Indels 8; Gaps 3;

Qy      5 HILIVEDLVTRNTILKSIFAEQYDVFEATDGAEMHQILSEYDINLVIMDINLPKNGLL 64
Db      6 HILIVDDDDREIRELVGNLYKKNGLTATIVADGRQMFLEANTVDLIVLDIMPGDDGULL 65
Qy      65 LARELREQANVA--LMFLTGRDNEVDKILGLIGADDYITKFPNPRELTIRARNLLSRM 122
Db      66 LCRELVRGKHKATPVMLTARNDETDRIIGLEMGADDYLTKPFSARELLARINAVLRER 125
Qy      123 NLG---TVSEERRSVESYKFNWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPGKI 179
Db      126 MLPPLNLTVSESRLLI---GFGQWQLDTSARHLDDAGTVALSGAEYRLRLRVFLDHPORV 182
Qy      180 QSRAELKKMTGRELKPHDRTVDVTIRIRKHFEPTDPTPEIATIHGEYRF 232
Db      183 LSRDQLNLNTQGREADIDRSIDLLVSLRQLRGDADAREPEYIKTVSRSEGYVF 235
```

Search completed: February 11, 2006, 02:43:26

Job time : 45 secs

GenCore version 5.1.7

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2006, 02:35:03 ; Search time 254 Seconds
(without alignments)

661.086 Million cell updates/sec

Title: US-10-613-990A-32

Perfect score: 1230

Sequence: 1 MQTPHILIVEDLVTRNLK.....PEIIATHGEGYRFGDLED 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1230	100.0	238	1 ARCA_ECO57	P089q3 escherichia
2	1230	100.0	238	1 ARCA_ECOL6	P089q2 escherichia
3	1230	100.0	238	1 ARCA_ECOL1	P089q1 escherichia
4	1230	100.0	238	1 ARCA_SHIFL	P089q4 shigella fl
5	1227	99.8	238	2 Q57G13_SALCH	Q57g13 salmonella
6	1227	99.8	238	2 Q7CP63_SALTY	Q7cp63 salmonella
7	1227	99.8	238	2 Q8XFN2_SALTI	Q8xfn2 salmonella
8	1219	99.1	238	2 Q5PN99_SALPA	Q5pn99 salmonella
9	1145	93.1	238	2 Q66EU0_YERPS	Q66eu0 yersinia ps
10	1145	93.1	238	2 Q8ZIN7_YERPE	Q8zin7 yersinia pe
11	1144	93.0	238	2 Q6D0A6_ERWCT	Q6d0a6 erwinia car
12	1138	92.5	238	2 Q7N8Z7_PHOLL	Q7n8z7 photorhabd
13	1098	89.3	238	2 Q9KPG9_VIBCH	Q9kpg9 vibrio chol
14	1073	87.2	238	2 Q87SD5_VIBPA	Q87sd5 vibrio para
15	1064	86.5	207	2 Q5MDG2_ECOLI	Q5mdg2 escherichia
16	1061	86.3	206	2 Q5MDG1_ECOLI	Q5mdg1 escherichia
17	1060	86.2	239	2 Q6LJQ5_PHOPR	Q6ljq5 photobacter
18	1056	85.9	238	2 Q8DEN9_VIBVU	Q8den9 vibrio vuln
19	1056	85.9	238	2 Q7MNR9_VIBVU	Q7mnr9 vibrio vuln
20	1056	85.9	239	2 Q5E2Y1_VIBF1	Q5e2y1 vibrio fisc
21	1052	85.5	238	2 Q56B54_VIBVU	Q56b54 vibrio vuln
22	1051	85.4	204	2 Q5MDP8_ECOLI	Q5mdf8 escherichia
23	1043	84.8	203	2 Q5MDF9_ECOLI	Q5mdf9 escherichia
24	1012.5	82.3	236	2 Q9CP42_PASMU	Q9cp42 pasteurella
25	1006	81.8	238	2 Q8EAB9_SHEON	Q8eab9 shewanella
26	987.5	80.3	236	2 Q65SE9_MANSM	Q65se9 manheimia
27	972.5	79.1	236	2 Q4QM24_HAB18	Q4qm24 haemophilus
28	968	78.7	188	2 Q9K366_ECOLI	Q9k366 escherichia
29	968	78.7	237	2 Q7VP30_HABDU	Q7vp30 haemophilus
30	964.5	78.4	236	1 ARCA_HAEIN	P44918 haemophilus
31	963	78.3	188	2 Q9KH76_ECOLI	Q9kh76 escherichia

32	941	76.5	183	2	Q5IP84_ECOLI	Q5ip84 escherichia
33	941	76.5	183	2	Q5IP89_SHIFL	Q5ip89 shigella fl
34	941	76.5	183	2	Q5IP91_SHIDY	Q5ip91 shigella dy
35	941	76.5	183	2	Q5IP93_SHISO	Q5ip93 shigella so
36	936	76.1	182	2	Q5IP77_ECOLI	Q5ip77 escherichia
37	936	76.1	182	2	Q5IP98_SHISO	Q5ip98 shigella so
38	936	76.1	182	2	Q5IP99_SHIBO	Q5ip99 shigella bo
39	936	76.1	182	2	Q5IP95_SHIFL	Q5ip95 shigella fl
40	934	75.9	182	2	Q5IP88_SHIFL	Q5ip88 shigella fl
41	934	75.9	182	2	Q5IP94_SHIDY	Q5ip94 shigella dy
42	934	75.9	182	2	Q5IP81_ECOLI	Q5ip81 escherichia
43	930	75.6	181	2	Q5IP80_ECOLI	Q5ip80 escherichia
44	929	75.5	181	2	Q5IP90_SHIBO	Q5ip90 shigella bo
45	929	75.5	181	2	Q65C04_ECOLI	Q65c04 escherichia

ALIGNMENTS

RESULT 1

ARCA_ECO57
ID ARCA_ECO57 STANDARD; PRT; 238 AA.
AC P0A9Q3; P03026;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Aerobic respiration control protein arca.
GN Name=arca; OrderedLocusNames=z6004, EC65359;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
[2]
RC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=O157:H7 / Sakai / RMD 0509952 / EHEC;
RX MEDLINE=2115231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Member of the two-component regulatory system arcB/arcA.
CC Represses a wide variety of aerobic enzymes under anaerobic
CC conditions. It also may be involved in the osmoregulation of
CC envelope proteins. When activated by arcB, it negatively regulates
CC the expression of genes of aerobic function. Activates the
CC transcription of the plfB operon by binding to its promoter (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- PTM: Phosphorylated by arcB (By similarity).
CC -!- SIMILARITY: Contains 1 response regulatory domain.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

CC EMBL; AE005174; AAG59581.1; -; Genomic_DNA.
DR

```
DR EMBL; BA000007; BAB38782.1; -; Genomic_DNA.
DR PIR; A86140; A86140.
DR PIR; G91298; G91298.
DR HSP; P13792; IMVO.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF000072; Response_reg; 1.
DR Pfam; PF00486; Trans_reg_C; 1.
DR ProDom; PD000039; Response_reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.
KW Activator; Complete proteome; DNA-binding; Phosphorylation; Repressor;
KW Sensory transduction; Transcription; Transcription regulation;
KW Two-component regulatory system.
FT DOMAIN 1 118 Response_reg; 1.
FT MOD RES 54 54 4-aspartylphosphate (By similarity).
SQ SEQUENCE 238 AA; 27292 MW; 9316CF4DE8EABDE8 CRC64;

Query Match 100.0%; Score 1230; DB 1; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.6e-88;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
DB 1 MOTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
QY 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRELTIARNLLSR 120
DB 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRELTIARNLLSR 120
QY 121 TMNLGTVEERRSVESYKFNWELDINSRSLIGPDGEQYKLPSEFRAMLHFCENPGKIQ 180
DB 121 TMNLGTVEERRSVESYKFNWELDINSRSLIGPDGEQYKLPSEFRAMLHFCENPGKIQ 180
QY 181 SRAELLKMTGRELKPHDRTVDVTRIRKHFESTPTPEIIATIHGEGYFCGDLED 238
DB 181 SRAELLKMTGRELKPHDRTVDVTRIRKHFESTPTPEIIATIHGEGYFCGDLED 238

RESULT 2
ARCA_ECOLI
ID ARCA_ECOLI STANDARD; PRT; 238 AA.
AC POA9Q2; P03026;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Aerobic respiration control protein arca (Dye resistance protein).
GN Name=arca; Synonyms=cpxC, dye, fexA, msp, seg, sifa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -!- FUNCTION: Member of the two-component regulatory system arcaB/arca.
CC Represents a wide variety of aerobic enzymes under anaerobic
CC conditions. It also may be involved in the osmoregulation of
CC envelope proteins. When activated by arcaB, it negatively regulates
CC the expression of genes of aerobic function. Activates the
CC transcription of the pIB operon by binding to its promoter (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- PTM: Phosphorylated by arcaB (By similarity).
```

```
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC -!- CAUTION: There are two genes termed arca in strain O6 of E.coli,
CC one refers to an arginine deiminase and the other to a two-
CC component regulator.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE016772; AAN83908.1; -; Genomic_DNA.
CC HSP; P13792; IMVO.
CC InterPro; IPR001789; Response_reg.
CC InterPro; IPR001867; Trans_reg_C.
CC Pfam; PF000072; Response_reg; 1.
CC Pfam; PF00486; Trans_reg_C; 1.
CC ProDom; PD000039; Response_reg; 1.
CC ProDom; PD000329; Trans_reg_C; 1.
CC SMART; SM00448; REC; 1.
CC PROSITE; PS0110; RESPONSE_REGULATORY; 1.
KW Activator; Complete proteome; DNA-binding; Phosphorylation; Repressor;
KW Sensory transduction; Transcription; Transcription regulation;
KW Two-component regulatory system.
FT DOMAIN 1 118 Response_reg; 1.
FT MOD RES 54 54 4-aspartylphosphate (By similarity).
SQ SEQUENCE 238 AA; 27292 MW; 9316CF4DE8EABDE8 CRC64;

Query Match 100.0%; Score 1230; DB 1; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.6e-88;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
DB 1 MOTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
QY 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRELTIARNLLSR 120
DB 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRELTIARNLLSR 120
QY 121 TMNLGTVEERRSVESYKFNWELDINSRSLIGPDGEQYKLPSEFRAMLHFCENPGKIQ 180
DB 121 TMNLGTVEERRSVESYKFNWELDINSRSLIGPDGEQYKLPSEFRAMLHFCENPGKIQ 180
QY 181 SRAELLKMTGRELKPHDRTVDVTRIRKHFESTPTPEIIATIHGEGYFCGDLED 238
DB 181 SRAELLKMTGRELKPHDRTVDVTRIRKHFESTPTPEIIATIHGEGYFCGDLED 238

RESULT 3
ARCA_ECOLI
ID ARCA_ECOLI STANDARD; PRT; 238 AA.
AC POA9Q1; P03026;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Aerobic respiration control protein arca (Dye resistance protein).
GN Name=arca; Synonyms=cpxC, dye, fexA, msp, seg, sifa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=85157583; PubMed=2984198;
RA Drury L.S., Buxton R.S.;
RT "DNA sequence analysis of the dye gene of Escherichia coli reveals
RT amino acid homology between the dye and OmpR proteins.";
RL J. Biol. Chem. 260:4236-4242(1985).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K12 / MG1655;
```


RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the *Escherichia coli* genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes";
RL Nucleic Acids Res. 23:2105-2119(1995).
[3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of *Escherichia coli* K-12";
RL Science 277:1453-1474(1997).
[4]
RP NUCLEOTIDE SEQUENCE OF 1-23.
RC STRAIN=K12;
RA Park S.J., Gunsalus R.P.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
[5]
RP PROTEIN SEQUENCE OF 1-12.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of *Escherichia coli* K-12";
RL Electrophoresis 18:1259-1313(1997).
[6]
RP PROTEIN SEQUENCE OF 40-45 AND 123-128, AND CHARACTERIZATION.
RX MEDLINE=96037802; PubMed=7565118;
RA Drapal N., Sawers G.;
RT "Purification of ArcA and analysis of its specific interaction with
RT the pfl promoter-regulatory region";
RL Mol. Microbiol. 16:597-607(1995).
CC -1- FUNCTION: Member of the two-component regulatory system arcB/arcA.
CC Represents a wide variety of aerobic enzymes under anaerobic
CC conditions. Controls the resistance of *E. coli* to dyes; required
CC for expression of the alkaline phosphatase and sex factor P genes;
CC It also may be involved in the osmoregulation of envelope
CC proteins. When activated by arcB, it negatively regulates the
CC expression of genes of aerobic function. Activates the
CC transcription of the pflB operon by binding to its promoter.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- PTM: Phosphorylated by arcB.
CC -1- SIMILARITY: Contains 1 response regulatory domain.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; M10044; AAA23718.1; -; Genomic DNA.
CC EMBL; U14003; AAA97297.1; -; Genomic DNA.
CC EMBL; U00096; AAC77354.1; -; Genomic DNA.
CC EMBL; L34010; AAC23476.1; -; Genomic DNA.
CC PIR; A03561; JYECR.
CC HSP; F13792; IMVO.
CC SWISS-2DPAGE; P0A9Q1; COLI.
CC EcoGene; EG10061; arcA.
CC InterPro; IPR001789; Response_reg.
CC Pfam; PF00072; Response_reg; 1.
CC Pfam; PF00486; Trans_reg_C; 1.
CC ProDom; PD000039; Response_reg; 1.
CC ProDom; PD000329; Trans_reg_C; 1.
CC SMART; SM00448; REC; 1.
CC PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Activator; Complete proteome; Direct protein sequencing; DNA-binding;
Phosphorylation; Repressor; Sensory transduction; Transcription;

KW Transcription regulation; Two-component regulatory system.
FT DOMAIN 1 118 Response regulatory.
FT MOD_RES 54 58 4-aspartylphosphate (By similarity).
SQ SEQUENCE 238 AA; 27292 MW; 9316CF4DE8ABDE8 CRC64;
Query Match 100.0%; Score 1230; DB 1; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.6e-88;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MQCPHILIVDELVTRNTLSKIFEAEGYDYFEATDGAEMHQILSEYDINLVIMDINLPKG 60
Db 1 MQCPHILIVDELVTRNTLSKIFEAEGYDYFEATDGAEMHQILSEYDINLVIMDINLPKG 60
Qy 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKFPNPRELTIRARNLSR 120
Db 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKFPNPRELTIRARNLSR 120
Qy 121 TMNLGTVSEERRSVESYKFNQWELDINSRLIGPDGEQYKLPSEPRAMLHFCENPGKI 180
Db 121 TMNLGTVSEERRSVESYKFNQWELDINSRLIGPDGEQYKLPSEPRAMLHFCENPGKI 180
Qy 181 SRAELKKMTGRELKPDRVTVDVTRIRKHFSTPTPELIATIHGEGYRFGCDLED 238
Db 181 SRAELKKMTGRELKPDRVTVDVTRIRKHFSTPTPELIATIHGEGYRFGCDLED 238
RESULT 4
ID_ARCA_SHIFL STANDARD; PRT; 238 AA.
AC P0A9Q1; F03026;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Aerobic respiration control protein arcA.
GN Name=arcA; OrderedLocusNames=SF4433, S4704;
OS *Shigella flexneri*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; *Shigella*.
OX NCBI_TaxID=623;
[1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
RT through comparison with genomes of *Escherichia coli* K12 and O157";
RL Nucleic Acids Res. 30:4432-4441(2002).
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=2259274; PubMed=12704152;
RA DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of *Shigella*
RT *flexneri* serotype 2a strain 2457T";
RL Infect. Immun. 71:2775-2786(2003).
CC -1- FUNCTION: Member of the two-component regulatory system arcB/arcA.
CC Represents a wide variety of aerobic enzymes under anaerobic
CC conditions. It also may be involved in the osmoregulation of
CC envelope proteins. When activated by arcB, it negatively regulates
CC the expression of genes of aerobic function. Activates the
CC transcription of the pflB operon by binding to its promoter (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- PTM: Phosphorylated by arcB (By similarity).
CC -1- SIMILARITY: Contains 1 response regulatory domain.

```

CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE005674; AAN45846.1; ALT INIT; Genomic_DNA.
CC EMBL; AE016993; AAP19620.1; -; Genomic_DNA.
CC HSP; P13792; IMVO.
CC InterPro; IPR001789; Response reg.
CC InterPro; IPR001867; Trans reg C.
CC Pfam; PF00072; Response reg; 1.
CC Pfam; PF00486; Trans_reg_C; 1.
CC ProDom; PD000039; Response_reg; 1.
CC ProDom; PD000329; Trans_reg_C; 1.
CC SMART; SM00448; REC; 1.
CC PROSITE; PS0110; RESPONSE REGULATORY; 1.
CC Activator; Complete proteome; DNA-binding; Phosphorylation; Repressor;
CC Sensory transduction; Transcription; Transcription regulation;
CC Two-component regulatory system.
CC DOMAIN 1 118 Response regulatory.
FT MOD RES 54 54 4-aspartylphosphate (By similarity).
SQ SEQUENCE 238 AA; 27292 MW; 9316CF4DE8EABD8 CRC64;

Query Match 100.0%; Score 1230; DB 1; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.6e-88;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQLLSYDINLVIMDINLPKG 60
DB 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQLLSYDINLVIMDINLPKG 60

QY 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIRARNLLSR 120
DB 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIRARNLLSR 120

QY 121 TNNLGVTSSEERSVSYKFNGLWELINSRLSGDGEQYKLPFRSEFRAMLHFCENPGKIQ 180
DB 121 TNNLGVTSSEERSVSYKFNGLWELINSRLSGDGEQYKLPFRSEFRAMLHFCENPGKIQ 180

QY 181 SRAELLKMTGRELKPHDRTVDVTRIRKHFESTPDTPETIIATHGEGYRFGCDLED 238
DB 181 SRAELLKMTGRELKPHDRTVDVTRIRKHFESTPDTPETIIATHGEGYRFGCDLED 238

RESULT 5
ID Q57G13 SALCH PRELIMINARY; PRT; 238 AA.
AC Q57G13
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE Response regulator (OmpR family), in two-component regulatory system
DE with ArcB (Or CpxA), regulates genes in aerobic pathways.
GN Name-arcA; OrderedLocusNames=SC4443;
OS Salmonella choleraesuis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=591;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-867;
RX PubMed=15781495; DOI=10.1093/nar/gki297;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
RA Wang H.-S., Lee Y.-S.;
RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
RT highly invasive and resistant zoonotic pathogen.";
RL Nucleic Acids Res. 33:1690-1698(2005).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
DR EMBL; AE017220; AAX68349.1; -; Genomic_DNA.
DR SMR; Q57G13; 1-122.
DR GO; GO:0003677; F:DNA binding; IEA.

```

```

DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000150; P:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR001867; Trans_reg_C.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00072; Response_reg; 1.
DR Pfam; PF00486; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS0110; RESPONSE REGULATORY; 1.
DR Complete proteome; DNA-binding; Sensory transduction; Transcription;
DR Transcription regulation; Two-component regulatory system.
SQ SEQUENCE 238 AA; 27291 MW; 930B0F4DE8EABD8 CRC64;

Query Match 99.8%; Score 1227; DB 2; Length 238;
Best Local Similarity 99.6%; Pred. No. 6.2e-86;
Matches 237; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQLLSYDINLVIMDINLPKG 60
DB 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQLLSYDINLVIMDINLPKG 60

QY 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIRARNLLSR 120
DB 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIRARNLLSR 120

QY 121 TNNLGVTSSEERSVSYKFNGLWELINSRLSGDGEQYKLPFRSEFRAMLHFCENPGKIQ 180
DB 121 TNNLGVTSSEERSVSYKFNGLWELINSRLSGDGEQYKLPFRSEFRAMLHFCENPGKIQ 180

QY 181 SRAELLKMTGRELKPHDRTVDVTRIRKHFESTPDTPETIIATHGEGYRFGCDLED 238
DB 181 SRAELLKMTGRELKPHDRTVDVTRIRKHFESTPDTPETIIATHGEGYRFGCDLED 238

RESULT 6
ID Q7CP63 SALTY PRELIMINARY; PRT; 238 AA.
AC Q7CP63
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Response regulator (OmpR family), in two-component regulatory system
DE with ArcB (Or CpxA).
GN Name-arcA; OrderedLocusNames=STM4598;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP NUCLEOTIDE SEQUENCE
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
DR EMBL; AE008916; AAL23413.1; -; Genomic_DNA.
DR HSP; P08402; 1B00.
DR SMR; Q7CP63; 1-122.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0000156; P:two-component response regulator activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF00072; Response_reg; 1.
DR Pfam; PF00486; Trans_reg_C; 1.

```

```
DR ProDom; PD000039; Response_reg; 1.  
DR SMART; SM00448; REC; 1.  
DR PROSITE; PS01110; RESPONSE REGULATORY; 1.  
KW Complete proteome; DNA-binding; Sensory transduction; Transcription;  
Transcription regulation; Two-component regulatory system.  
SQ SEQUENCE 238 AA; 27291 MW; 93080F4D8EABD8 CRC64;  
  
Query Match 99.8%; Score 1227; DB 2; Length 238;  
Best Local Similarity 99.6%; Pred. No. 6.2e-88;  
Matches 237; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MQTPHILIVEDLVTRNLTKSIFEAAGYDVFEATDGAEMHQLSEYDINLVIMDINLPKG 60  
Db 1 MQTPHILIVEDLVTRNLTKSIFEAAGYDVFEATDGAEMHQLSEYDINLVIMDINLPKG 60  
  
Qy 61 NGLLARELRQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARNLLSR 120  
Db 61 NGLLARELRQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARNLLSR 120  
  
Qy 121 TNLGTVSEERSVESYKFNGLWELNDSRLIGPDGEQYKLPGRSEFRAMLHFCENPGKIQ 180  
Db 121 TNLGTVSEERSVESYKFNGLWELNDSRLIGPDGEQYKLPGRSEFRAMLHFCENPGKIQ 180  
  
Qy 181 SRAELKKMTGRELKHDPDRTVDVTRIRKHFESTPDTPEIATIHGSGYRFGDLED 238  
Db 181 SRAELKKMTGRELKHDPDRTVDVTRIRKHFESTPDTPEIATIHGSGYRFGDLED 238  
  
RESULT 7  
Q8XFN2 SALT PRELIMINARY; PRT; 238 AA.  
ID Q8XFN2; Q7ALM2;  
AC Q8XFN2; Q7ALM2;  
DT 01-MAR-2002 (TReMBLrel. 20, Created)  
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)  
DE Global response regulator.  
GN Name=arcA; OrderedLocusNames=STY4947, t4637;  
OS Salmonella typhi.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI_TaxID=601;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAINS=Ty2 / ATCC 700931;  
RX MEDLINE=22531367; PubMed=12644504;  
DOI=10.1128/JB.185.7.2330-2337.2003;  
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,  
Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;  
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2  
and CT18";  
RL J. Bacteriol. 185:2330-2337(2003).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAINS=CT18;  
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;  
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,  
Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,  
Whitehead S., Barrall B.G.;  
RT "Complete genome sequence of a multiple drug resistant Salmonella  
enterica serovar Typhi CT18";  
RL Nature 413:848-852(2001).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
DR EMBL; AE016849; AAO72066.1; -; Genomic_DNA.  
DR EMBL; AL627284; CAD03429.1; -; Genomic_DNA.  
DR HSP; F08402; I800.  
DR SMR; Q8XFN2; 1-122.  
DR GO; GO:0003677; F:DNA binding; IEA.
```

```
DR GO; GO:0000156; F:two-component response regulator activity; IEA.  
DR GO; GO:0007600; P:sensory perception; IEA.  
DR GO; GO:0006350; P:transcription; IEA.  
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.  
DR InterPro; IPR001789; Response_reg.  
DR InterPro; IPR001867; Trans_reg_C.  
DR InterPro; IPR011991; Wing_Hlx_DNA_bd.  
DR Pfam; PF00072; Response_reg; 1.  
DR Pfam; PF00486; Trans_reg_C; 1.  
DR ProDom; PD000039; Response_reg; 1.  
DR ProDom; PD000329; Trans_reg_C; 1.  
DR SMART; SM00448; REC; 1.  
DR PROSITE; PS01110; RESPONSE REGULATORY; 1.  
KW Complete proteome; DNA-binding; Sensory transduction; Transcription;  
Transcription regulation; Two-component regulatory system.  
SQ SEQUENCE 238 AA; 27291 MW; 93080F4D8EABD8 CRC64;  
  
Query Match 99.8%; Score 1227; DB 2; Length 238;  
Best Local Similarity 99.6%; Pred. No. 6.2e-88;  
Matches 237; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MQTPHILIVEDLVTRNLTKSIFEAAGYDVFEATDGAEMHQLSEYDINLVIMDINLPKG 60  
Db 1 MQTPHILIVEDLVTRNLTKSIFEAAGYDVFEATDGAEMHQLSEYDINLVIMDINLPKG 60  
  
Qy 61 NGLLARELRQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARNLLSR 120  
Db 61 NGLLARELRQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARNLLSR 120  
  
Qy 121 TNLGTVSEERSVESYKFNGLWELNDSRLIGPDGEQYKLPGRSEFRAMLHFCENPGKIQ 180  
Db 121 TNLGTVSEERSVESYKFNGLWELNDSRLIGPDGEQYKLPGRSEFRAMLHFCENPGKIQ 180  
  
Qy 181 SRAELKKMTGRELKHDPDRTVDVTRIRKHFESTPDTPEIATIHGSGYRFGDLED 238  
Db 181 SRAELKKMTGRELKHDPDRTVDVTRIRKHFESTPDTPEIATIHGSGYRFGDLED 238  
  
RESULT 8  
Q5PN99 SALPA PRELIMINARY; PRT; 238 AA.  
ID Q5PN99; SALPA PRELIMINARY;  
AC Q5PN99;  
DT 01-FEB-2005 (TReMBLrel. 29, Created)  
DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)  
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)  
DE Global response regulator.  
GN Name=arcA; OrderedLocusNames=SPA4408;  
OS Salmonella paratyphi-a.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI_TaxID=54388;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=ATCC 9150 / SARB42;  
RX PubMed=15531882; DOI=10.1038/ng1470;  
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,  
Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,  
Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,  
Kohlberg S., Strong C., Du P., Carter J., Kremizki C., Layman D.,  
Leonhardt S., Sun H., Fulton L., Nash W., Miner T., Minx P.,  
Delehaanty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,  
Speth J., Wilson R.K.;  
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-  
restricted serovars of Salmonella enterica that cause typhoid";  
RL Nat. Genet. 36:1268-1274(2004).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
DR EMBL; CP000026; AA080130.1; -; Genomic_DNA.  
DR SMR; Q5PN99; 1-122.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0000156; F:two-component response regulator activity; IEA.  
DR GO; GO:0007600; P:sensory perception; IEA.  
DR GO; GO:0006350; P:transcription; IEA.  
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
```

```

DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF000072; Response_reg; 1.
DR Pfam; PF00486; Trans_reg_C; 1.
DR ProDom; PD000039; Response_reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Complete proteome; DNA-binding; Sensory transduction; Transcription;
KW Transcription regulation; Two-component regulatory system.
SQ SEQUENCE 238 AA; 27261 MW; 9309DD9D8EABEEB CRC64;

Query Match 99.1%; Score 1219; DB 2; Length 238;
Best Local Similarity 99.2%; Pred. No. 2.6e-87;
Matches 236; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MQTPHILVDELVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
Db 1 MQTPHILVDELVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
Qy 61 NGLLARELREQANVALMFLTGKNEVDKILGLEIGADDYITKPNPRELTIRARNLLSR 120
Db 61 NGLLARELREQANVALMFLTGKNEVDKILGLEIGADDYITKPNPRELTIRARNLLSR 120
Qy 121 TMNLGTVSEERSVESYKFNWELNDSRLSGPDGEQYKLPKPRSEFRAMLHFCENPGKIQ 180
Db 121 TMNLGTVSEERSVESYKFNWELNDSRLSGPDGEQYKLPKPRSEFRAMLHFCENPGKIQ 180
Qy 181 SRAELLKMTGRELKPHDRTVDVTRIRKHFESTPTDPEIIATIHGEGYRFGCDLED 238
Db 181 SRAELLKMTGRELKPHDRTVDVTRIRKHFESTPTDPEIIATIHGEGYRFGCDLED 238

RESULT 10
Q8ZIN7 YERPE PRELIMINARY; PRT; 238 AA.
AC Q8ZIN7; Q74Q02; Q7CG70; 20, Created
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Aerobic respiration control protein (Negative response regulator of
DE genes in aerobic pathways, sensors, ArcB and CpxA).
GN Name=arcA; Synonyms=cpxC, dye, fexA, map, seg, sfrA;
GN OrderedLocNames=yf3725, YPO0458, y3721;
OS Versinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CO-92 / Biovar Orientalis;
RC MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Church C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Versinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RC MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Versinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=91001;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Versinia pestis strain 91001, an isolate
RT avirulent to humans.";

```



```
DR InterPro; IPR001789; Response_reg;
DR InterPro; IPR001867; Trans_reg_C;
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00072; Response_reg; 1.
DR Pfam; PF00486; Trans_reg_C; 1.
DR ProDom; PD000039; Response_reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Complete proteome; DNA-binding; Sensory transduction; Transcription;
KW Transcription regulation; Two-component regulatory system.
SQ SEQUENCE 238 AA; 27296 MW; 46DF244FC0C21D5D CRC64;

Query Match 92.5%; Score 1138; DB 2; Length 238;
Best Local Similarity 93.2%; Pred. No. 5.8e-81;
Matches 221; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MOTPHILIVEDLVRNTLKSIFFAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
Db 1 MOTPHILIVEDLVRNTLKSIFFAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60

Qy 61 NGLLARELREQANVALMFLTRGRNEVDKILGLEIGADDYTKPNPRELTIRARNLLSR 120
Db 61 NGLLARELREQANVALMFLTRGRNEVDKILGLEIGADDYTKPNPRELTIRARNLLSR 120

Qy 121 TMLNLTGVSEERSVESYKFNWELDINSRLSPAGEYKLPKPRSEFRAMLHFCENPGKIQ 180
Db 121 TMLNSVSEERSQVESYKFNWELDINSRLSPAGEYKLPKPRSEFRAMLHFCENPGKIQ 180

Qy 181 SRAELLKMTGRELKPHDRVTVDVTRIRKHFESTPTDPTPEIIATIHGEGYRFGCDLE 237
Db 181 TRADLLKMTGRELKPHDRVTVDVTRIRKHFESTPTDPTPEIIATIHGEGYRFGCDLE 237

RESULT 13
Q9KJP9 VIBCH
ID Q9KJP9 VIBCH PRELIMINARY; PRT; 238 AA.
AC Q9KJP9
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Aerobic respiration control protein FexA.
GN OrderedLocusNames=VC2368;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Uitterback T.R., Fleischmann R.D.,
RA Niernan W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.";
RL Nature 406:477-483 (2000).
DR EMBL; AE004307; AAF95511.1; -; Genomic_DNA.
DR FIC; H82084; H82084.
DR HSP; F08402; 1B00.
DR SMR; Q9KJP9; 1-122.
DR TIGR; VC2368; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:000156; P:two-component response regulator activity; IEA.
DR GO; GO:000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR001867; Trans_reg_C.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00072; Response_reg; 1.
DR Pfam; PF00486; Trans_reg_C; 1.
```

```
DR ProDom; PD000039; Response_reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Complete proteome.
SQ SEQUENCE 238 AA; 27119 MW; 6F59FBFD5C4C2540 CRC64;

Query Match 89.3%; Score 1098; DB 2; Length 238;
Best Local Similarity 88.2%; Pred. No. 7.8e-76;
Matches 210; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

Qy 1 MOTPHILIVEDLVRNTLKSIFFAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
Db 1 MOTPHILIVEDLVRNTLKSIFFAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60

Qy 61 NGLLARELREQANVALMFLTRGRNEVDKILGLEIGADDYTKPNPRELTIRARNLLSR 120
Db 61 NGLLARELREQADVALMFLTRGRNEVDKILGLEIGADDYTKPNPRELTIRARNLLSR 120

Qy 121 TMLNLTGVSEERSVESYKFNWELDINSRLSPAGEYKLPKPRSEFRAMLHFCENPGKIQ 180
Db 121 TMLNLTGVSEERSVESYKFNWELDINSRLSPAGEYKLPKPRSEFRAMLHFCENPGKIQ 180

Qy 181 SRAELLKMTGRELKPHDRVTVDVTRIRKHFESTPTDPTPEIIATIHGEGYRFGCDLE 238
Db 181 TRADLLKMTGRELKPHDRVTVDVTRIRKHFESVSGTPEIIATIHGEGYRFGCDLE 238

RESULT 14
Q87SD5 VIBPA
ID Q87SD5 VIBPA PRELIMINARY; PRT; 238 AA.
AC Q87SD5
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Aerobic respiration control protein FexA.
GN OrderedLocusNames=VP0489;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
distinct from that of V. cholerae.";
RL Lancet 361:743-749 (2003).
DR EMBL; BA000031; BACS8752.1; -; Genomic_DNA.
DR HSP; F08402; 1B00.
DR SMR; Q87SD5; 1-122.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:000156; P:two-component response regulator activity; IEA.
DR GO; GO:000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR001867; Trans_reg_C.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00072; Response_reg; 1.
DR Pfam; PF00486; Trans_reg_C; 1.
DR ProDom; PD000039; Response_reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Complete proteome.
SQ SEQUENCE 238 AA; 26936 MW; 14429758CC899F78 CRC64;

Query Match 87.2%; Score 1073; DB 2; Length 238;
Best Local Similarity 86.6%; Pred. No. 7.1e-76;
Matches 206; Conservative 13; Mismatches 19; Indels 0; Gaps 0;
```

Qy 1 MOTPHILIVEDELVTRNTLKSIFEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPK 60
Db 1 MOTPHILIVEDELVTRNTLKSIFEAGYDVFEATDGAEMHQVLSNINLVIMDINLPK 60
Qy 61 NGLLIARELREQANVALMFLTGRDNEVDKILGLEIGADDYTKPNPRELTIRARNLLSR 120
Db 61 NGLLIARELREQANVALMFLTGRDNEVDKILGLEIGADDYTKPNPRELTIRARNLLSR 120
Qy 121 TMTNLGTVSEERSVSEYKFNWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPGKI 180
Db 121 TMTNLGTVSEERSVSEYKFNWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPGKI 180
Qy 181 SRAELLLKMTGRELKPHDRTVDVTRIRIRKHFESTPTDTPETIATIHGEGYRFGDLED 238
Db 181 SRAELLLKMTGRELKPHDRTVDVTRIRIRKHFESTPTDTPETIATIHGEGYRFGDLED 238

RESULT 15

Q5MDG2_ECOLI PRELIMINARY; PRT; 207 AA.
AC Q5MDG2;
DT 01-FEB-2005 (TremBLrel. 29, Created)
DT 01-FEB-2005 (TremBLrel. 29, Last sequence update)
DE Aerobic respirator control protein A (Fragment).
GN Name=arx4;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_taxid=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=044;
RX PubMed=15814998; DOI=10.1128/JCM.43.4.1776-1781.2005;
RA Nemoy L.L., Kotetishvili M., Tigmo J., Keefer-Norris A., Harris A.D.,
RA Perencevich E.N., Johnson J.A., Torpey D., Sulakvelidze A.,
RA Morris J.G. Jr., Stine O.C.;
RT "Multilocus sequence typing versus pulsed-field gel electrophoresis
RT for characterization of extended-spectrum beta-lactamase-producing
RT Escherichia coli isolates";
RL J. Clin. Microbiol. 43:1776-1781(2005).
DR EMBL; AY832936; AAV91697.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF00072; Response_reg; 1.
DR Pfam; PF00486; Trans_reg_C; 1.
DR ProDom; PD000039; Response_reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS01110; RESPONSE_REGULATORY; 1.
KW DNA-binding; Sensory transduction; Transcription;
KW Transcription regulation.
FT NON_TER 1
FT NON_TER 207
SQ SEQUENCE 207 AA; 23745 MW; D821137D4D703F9D CRC64;

Query Match 86.5%; Score 1064; DB 2; Length 207;
Best Local Similarity 100.0%; Pred. No. 3e-75;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 KSIPEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPKNGLLIARELREQANVALMP 79
Db 1 KSIPEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPKNGLLIARELREQANVALMP 60
Qy 80 LTGRDNEVDKILGLEIGADDYTKPNPRELTIRARNLLSRMTNLGTVSEERSVSEYKF 139
Db 61 LTGRDNEVDKILGLEIGADDYTKPNPRELTIRARNLLSRMTNLGTVSEERSVSEYKF 120

Qy 140 NGWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPGKIQSRAELLLKMTGRELKPHDR 199
Db 121 NGWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPGKIQSRAELLLKMTGRELKPHDR 180
Qy 200 TVDVTIRIRKHFESTPTDTPETIATIH 226
Db 181 TVDVTIRIRKHFESTPTDTPETIATIH 207

Search completed: February 11, 2006, 02:42:38
Job time : 256 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2006, 02:42:54 ; Search time 48 Seconds
(without alignments)
409.934 Million cell updates/sec

Title: US-10-613-990A-32
Perfect score: 1230
Sequence: 1 MQTPHILIVEDLVTRNTLK.....PEIIATIHGEGYRFGCDLED 238

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents_AA.*
1: /cgm2_6/ptodata/1/iaa/5-COMB.pep.*
2: /cgm2_6/ptodata/1/iaa/6-COMB.pep.*
3: /cgm2_6/ptodata/1/iaa/H-COMB.pep.*
4: /cgm2_6/ptodata/1/iaa/PCUTUS-COMB.pep.*
5: /cgm2_6/ptodata/1/iaa/RE-COMB.pep.*
6: /cgm2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1197	97.3	241	2	US-09-489-039A-8983 Sequence 8983, Ap
2	1133	92.1	241	2	US-09-543-681A-6205 Sequence 6205, Ap
3	458	37.2	278	2	US-09-252-991A-32988 Sequence 32988, A
4	400	32.5	721	2	US-09-252-991A-28293 Sequence 28293, A
5	384.5	31.3	249	2	US-09-543-681A-6721 Sequence 6721, Ap
6	376	30.6	246	2	US-09-107-532A-4135 Sequence 4135, Ap
7	375	30.5	233	2	US-09-082-077-4 Sequence 4, Appl
8	375	30.5	233	2	US-09-710-279-480 Sequence 480, App
9	375	30.5	233	2	US-09-925-637-46 Sequence 46, Appl
10	375	30.5	234	2	US-09-134-001C-3063 Sequence 3063, Ap
11	375	30.5	239	2	US-08-956-171E-5239 Sequence 5239, Ap
12	375	30.5	239	2	US-08-781-986A-5239 Sequence 5239, Ap
13	374	30.4	294	2	US-09-489-039A-9766 Sequence 9766, Ap
14	373.5	30.4	260	2	US-09-489-039A-10893 Sequence 10893, A
15	372.5	30.3	239	1	US-08-286-819A-44 Sequence 44, Appl
16	372.5	30.3	239	2	US-08-980-357-44 Sequence 44, Appl
17	372.5	30.3	239	2	US-09-357-375-44 Sequence 44, Appl
18	370	30.1	276	2	US-09-252-991A-21689 Sequence 21689, A
19	365	29.7	234	1	US-08-850-118-2 Sequence 2, Appl
20	365	29.7	234	1	US-09-008-253-2 Sequence 2, Appl
21	365	29.7	234	2	US-09-093-335-2 Sequence 2, Appl
22	365	29.7	234	2	US-09-583-110-4584 Sequence 4584, Ap
23	365	29.7	252	2	US-09-107-433-3288 Sequence 3288, Ap
24	362	28.4	246	2	US-09-134-001C-3175 Sequence 3175, Ap
25	349.5	28.4	265	2	US-09-328-352-6073 Sequence 6073, Ap
26	346.5	28.2	234	2	US-09-634-238-220 Sequence 220, App
27	345.5	28.1	236	2	US-09-543-681A-8023 Sequence 8023, Ap

ALIGNMENTS

RESULT 1

US-09-489-039A-8983
; Sequence 8983, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIORITY FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8983
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-8983

Query Match	97.3%	Score 1197;	DB 2;	Length 241;
Best Local Similarity	97.1%	Pred. No. 3.2e-125;		
Matches	231;	Conservative	5;	Mismatches 2;
		Indels	0;	Gaps 0;
Qy	1	MQTPHILIVEDLVTRNTLKSI	FEAEGYDVPEATDGAEMHQILSEYDINLVIMDINLPCK	60
Db	4	MQTPHILIVEDLVTRNTLKSI	FEAEGYDVPEATDGAEMHQILSENDINLVIMDINLPCK	63
Qy	61	NGLLARELREQANVALMFLTG	RNEVDKIILGLEIGADDDYITKPNPRELTIARNLISR	120
Db	64	NGLLARELREQADVALMFLTG	RNEVDKIILGLEIGADDDYITKPNPRELTIARNLISR	123
Qy	121	TMNLTGVSEERSVESYKFN	GWELDINSRSLIGPDGEQYKLPSEFRAMLHFCENPGKIQ	180
Db	124	TMNLTGVSEERSVESYKFN	GWELDINSRSLVSPNGEQYKLPSEFRAMLHFCENPGKIQ	183
Qy	181	SRAELLKQMTGRELKPHDR	TVDVTIRIRKHFESTPDTPPEIIATIHGEGYRFGCDLED	238
Db	184	SRAELLKQMTGRELKPHDR	TVDVTIRIRKHFESTPDTPPEIIATIHGEGYRFGCDLQE	241

RESULT 2

US-09-543-681A-6205
; Sequence 6205, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001


```
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 480
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-480

Query Match          30.5%; Score 375; DB 2; Length 233;
Best Local Similarity 35.6%; Pred. No. 1.8e-33;
Matches 83; Conservative 52; Mismatches 80; Indels 18; Gaps 3;

QY 6 ILIYDELVTRNTLKSIFEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPGKNGLL 65
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
5 VVVVDDEKPIADILEFNLKKEGYDVYCAVDGNDVLDLIEBEPDIVLLDIMPGRDGMV 64

QY 66 ARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARLLSRTMNLG 125
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
65 CREVRKKYEMPIMLTAKDSEIDKVLGELGADDYVTKPFSTRELIAKVKANLRR----- 119

QY 126 TVSEERRSVESYKFGWELDINSRSL-IGPD-----GEQYKLPSEFRAMLHFCENPG 177
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
120 -----HYSQPAQEVSGATNEITIKDIVYPPDAYSIKKGGEDIELTHREFELPHYLSKMG 174

QY 178 KIQSRAELLKMTGRELKPHDRVTVDVTRIRKHFESTPDTPPEIATIHGEY 230
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
175 QVMTREHLQTVMGYDYGVDVTRVTRIRREKIEDDPSPHPEYIVTRRGVGY 227

RESULT 9
US-09-925-637-46
; Sequence 46, Application US/09925637
; Patent No. 6833253
; GENERAL INFORMATION:
; APPLICANT: Choi
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925,637
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-637-46

Query Match          30.5%; Score 375; DB 2; Length 233;
Best Local Similarity 35.3%; Pred. No. 1.8e-33;
Matches 84; Conservative 50; Mismatches 76; Indels 28; Gaps 4;

QY 6 ILIYDELVTRNTLKSIFEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPGKNGLL 65
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
5 VVVVDDEKPIADILEFNLKKEGYDVYCAVDGNDVLDLIEBEPDIVLLDIMPGRDGMV 64

QY 66 ARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARLLSRTMNLG 120
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
65 CREVRKKYEMPIMLTAKDSEIDKVLGELGADDYVTKPFSTRELIAKVKANLRRHSQP 124
```

```
QY 121 TMNLGTVSEERRSVESYKFGWELDINSRSL-IGPD-----GEQYKLPSEFRAMLHF 172
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
125 AQDTGNVTNE-----ITIKDIVYPPDAYSIKKGGEDIELTHREFELPHYL 169

QY 173 CENPGKIQSRAELLKMTGRELKPHDRVTVDVTRIRKHFESTPDTPPEIATIHGEY 230
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
170 SKHMGVMTREHLQTVMGYDYGVDVTRVTRIRREKIEDDPSPHPEYIVTRRGVGY 227

RESULT 10
US-09-134-001C-3063
; Sequence 3063, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3063
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3063

Query Match          30.5%; Score 375; DB 2; Length 234;
Best Local Similarity 35.6%; Pred. No. 1.8e-33;
Matches 83; Conservative 52; Mismatches 80; Indels 18; Gaps 3;

QY 6 ILIYDELVTRNTLKSIFEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPGKNGLL 65
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
6 VVVVDDEKPIADILEFNLKKEGYDVYCAVDGNDVLDLIEBEPDIVLLDIMPGRDGMV 65

QY 66 ARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARLLSRTMNLG 125
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
66 CREVRKKYEMPIMLTAKDSEIDKVLGELGADDYVTKPFSTRELIAKVKANLRR----- 120

QY 126 TVSEERRSVESYKFGWELDINSRSL-IGPD-----GEQYKLPSEFRAMLHFCENPG 177
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 -----HYSQPAQEVSGATNEITIKDIVYPPDAYSIKKGGEDIELTHREFELPHYLSKMG 175

QY 178 KIQSRAELLKMTGRELKPHDRVTVDVTRIRKHFESTPDTPPEIATIHGEY 230
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
176 QVMTREHLQTVMGYDYGVDVTRVTRIRREKIEDDPSPHPEYIVTRRGVGY 228

RESULT 11
US-08-956-171E-5239
; Sequence 5239, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; GIL H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 5239:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5239:
US-08-956-171E-5239

Query Match 30.5%; Score 375; DB 2; Length 239;
Best Local Similarity 35.3%; Pred. No. 1.9e-33;
Matches 84; Conservative 50; Mismatches 76; Indels 28; Gaps 4;
QY 6 ILIIVEDLVTRNTLKSIFEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPKNGLL 65
DB 11 VVVVDEKPIADILEFNLKKEGYDVYCAVDGNDVLDIYEEPDVILDDIMLPGRDGM 70
QY 66 ARELREQANVALMFLTGDRNEVDKILGLEIGADDYITKPFNPRLTIRARNLLSR- 120
DB 71 CREVKKYEMPIMLTAKDSEIDKVLGLELGGADYVTKPFSTRELIAKVKANLRRHYSQ 130
QY 121 TNLGTVSEERSVSEYKFNKGWELDINSRL-IGPD-----GEOYKLPSEFRAMLHF 172
DB 131 AODTGNVNE-----ITIKDIVPDAYSIKKRGSDIELTHREFELPHYL 175
QY 173 CENPGKIQSRAELKKMTGRELKPHDRTVDTTIRIRKHFSTPTDPTPEIIATIHGEY 230
DB 176 SKHMGQVMTREHLQTVMGYDFGVDVTVTIRLRKEIEDDPSPHPYIVTRRGVY 233

RESULT 12
US-08-781-986A-5239
Sequence 5239, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSES: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PPP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 5239:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-781-986A-5239

Query Match 30.5%; Score 375; DB 2; Length 239;
Best Local Similarity 35.3%; Pred. No. 1.9e-33;
Matches 84; Conservative 50; Mismatches 76; Indels 28; Gaps 4;
QY 6 ILIIVEDLVTRNTLKSIFEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPKNGLL 65
DB 11 VVVVDEKPIADILEFNLKKEGYDVYCAVDGNDVLDIYEEPDVILDDIMLPGRDGM 70
QY 66 ARELREQANVALMFLTGDRNEVDKILGLEIGADDYITKPFNPRLTIRARNLLSR- 120
DB 71 CREVKKYEMPIMLTAKDSEIDKVLGLELGGADYVTKPFSTRELIAKVKANLRRHYSQ 130
QY 121 TNLGTVSEERSVSEYKFNKGWELDINSRL-IGPD-----GEOYKLPSEFRAMLHF 172
DB 131 AODTGNVNE-----ITIKDIVPDAYSIKKRGSDIELTHREFELPHYL 175
QY 173 CENPGKIQSRAELKKMTGRELKPHDRTVDTTIRIRKHFSTPTDPTPEIIATIHGEY 230
DB 176 SKHMGQVMTREHLQTVMGYDFGVDVTVTIRLRKEIEDDPSPHPYIVTRRGVY 233

RESULT 13
US-09-489-039A-9766
Sequence 9766, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 9766
LENGTH: 294
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9766

Query Match 30.4%; Score 374; DB 2; Length 294;
Best Local Similarity 34.6%; Pred. No. 3.3e-33;
Matches 80; Conservative 50; Mismatches 93; Indels 8; Gaps 3;
QY 5 HILIVEDLVTRNTLKSIFEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPKNGLL 64
DB 55 HILVDDDDRIRELIVDYLEKSGYRASGAANGKAMWSVLKNHQIDLIIVLDIMMPGEDGLT 114
QY 65 LARELR--EQANVALMFLTGDRNEVDKILGLEIGADDYITKPFNPRLTIRARNLLSR 122
DB 115 LCRQLRANPOQDIPVLMLTARTDDSDRLGLEMGADDYLIKPFVARELLARIKALRRTR 174

```
QY 123 NLG---TVSEERRSVESYKFNKGWELDNRSRLIGPDGQYKLPSEFRAMLHFCENPGKI 179
Db 175 ALPPNQLQTEAGRLI---AFGWLMLDTARHLDDSGAIVASGAYEYLLRVFLDHPQV 231
QY 180 OSRAELKKMTGRELKPDRDVTVDVIRIRKHFESTPTDPTPEIATIHGEY 230
Db 232 LNRDQLNLITQGRDAELPERSIDLIVSLRQLRDLREDAEPAYIKTVRSEGY 282

RESULT 14
US-09-489-039A-10893
; Sequence 10893, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10893
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10893

Query Match 30.4%; Score 373.5; DB 2; Length 260;
Best Local Similarity 36.5%; Pred. No. 3.1e-33;
Matches 85; Conservative 47; Mismatches 94; Indels 7; Gaps 4;

QY 6 ILIIVDELVTRNTLKSIFEABGYDVFATDGAEMHQILSEYDINIVMDINLPGKNGLL 65
Db 28 ILVVDMMRLRALLERYLTQGFQVRSVANAEQMDRLLTRESFHLWLDLMLPGEDGLSI 87
QY 66 ARELREQAN-VALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIRARNLLSRWN- 123
Db 88 CRRLSQSNPFIIMVTAKGEEVDRIVGLGLEIGADDYITKPNPRELLARIARVLRQANE 147
QY 124 -LGTVSEERRSVESYKFNKGWELDNRSRLIGPDGQYKLPSEFRAMLHFCENPGKIQSR 182
Db 148 LPGAPSQEEAVTAFGKF---KLNLTGTFREDFRED-EPMLTSGEFAVLKALVSHPREPLSR 203
QY 183 AELLKMTGRELKPDRDVTVDVIRIRKHFESTPTDPTPEIATIHGEYRFGCD 235
Db 204 DKLMNLARGREYSAMERSIDVQISRLRRMVESDPAHPRYIQTVMGLGYVFPD 256

RESULT 15
US-08-286-819A-44
; Sequence 44, Application US/08286819A
; Patent No. 5871910
; GENERAL INFORMATION:
; APPLICANT: ARTHUR, MICHEL
; APPLICANT: DUKTA-MALEN, SYLVIE
; APPLICANT: MOLINAS, CATHERINE
; APPLICANT: COURVALIN, PATRICE
; TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
; TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPETIDES, IN PARTICULAR
; TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
; TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,819A
; FILING DATE: 05-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/917,146
; FILING DATE: 10-AUG-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR/91/00855
; FILING DATE: 29-OCT-1991
; APPLICATION NUMBER: FR 9013579
; FILING DATE: 31-OCT-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5871910man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-060-0 PCT
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
US-08-286-819A-44
```

```
Query Match 30.3%; Score 372.5; DB 1; Length 239;
Best Local Similarity 36.5%; Pred. No. 3.5e-33;
Matches 85; Conservative 47; Mismatches 94; Indels 7; Gaps 4;
```

```
QY 6 ILIIVDELVTRNTLKSIFEABGYDVFATDGAEMHQILSEYDINIVMDINLPGKNGLL 65
Db 7 ILVVDMMRLRALLERYLTQGFQVRSVANAEQMDRLLTRESFHLWLDLMLPGEDGLSI 66
```

```
QY 66 ARELREQAN-VALMFLTGRDNEVDKILGLEIGADDYITKPNPRELTIRARNLLSRWN- 123
Db 67 CRRLSQSNPFIIMVTAKGEEVDRIVGLGLEIGADDYITKPNPRELLARIARVLRQANE 126
```

```
QY 124 -LGTVSEERRSVESYKFNKGWELDNRSRLIGPDGQYKLPSEFRAMLHFCENPGKIQSR 182
Db 127 LPGAPSQEEAVTAFGKF---KLNLTGTFREDFRED-EPMLTSGEFAVLKALVSHPREPLSR 182
```

```
QY 183 AELLKMTGRELKPDRDVTVDVIRIRKHFESTPTDPTPEIATIHGEYRFGCD 235
Db 183 DKLMNLARGREYSAMERSIDVQISRLRRMVESDPAHPRYIQTVMGLGYVFPD 235
```

```
Search completed: February 11, 2006, 02:44:20
Job time : 49 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2006, 02:54:50 ; Search time 174 Seconds
(without alignments)
571.514 Million cell updates/sec

Title: US-10-613-990A-32
Perfect score: 1230
Sequence: 1 MQTPHILIVEDLVTRNTLK.....PEIIATIHGEGYRFGCDLED 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /cgm2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgm2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgm2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgm2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 5: /cgm2_6/ptodata/1/pubpaa/US108_PUBCOMB.pep.*
- 6: /cgm2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1230	100.0	238	4	US-10-613-990A-32
2	1145	93.1	238	4	US-10-613-990A-20
3	964.5	78.4	236	3	US-09-815-242-11143
4	964.5	78.4	236	4	US-10-282-122A-58341
5	778	63.3	176	5	US-10-450-763-59111
6	400	32.5	236	4	US-10-282-122A-74434
7	389.5	31.7	237	4	US-10-282-122A-60754
8	386.5	31.4	235	4	US-10-282-122A-46382
9	386.5	31.4	235	4	US-10-282-122A-72172
10	376	30.6	233	3	US-09-815-242-4913
11	375	30.6	234	3	US-09-815-242-10748
12	376	30.6	234	4	US-10-282-122A-42560
13	376	30.6	238	4	US-10-282-122A-71628
14	376	30.6	261	4	US-10-282-122A-60754
15	375	30.5	233	3	US-09-815-242-5676
16	375	30.5	233	3	US-09-925-637-46
17	375	30.5	233	4	US-10-084-205-46
18	375	30.5	233	4	US-10-712-713-46
19	375	30.5	233	5	US-10-923-429-46
20	375	30.5	234	4	US-10-724-972A-6908
21	375	30.5	235	3	US-09-815-242-12674
22	375	30.5	235	3	US-09-815-242-13000
23	375	30.5	235	4	US-10-282-122A-44277
24	375	30.5	235	5	US-10-857-625-768
25	375	30.5	239	2	US-08-781-986A-5239
26	375	30.5	239	4	US-10-282-122A-70821
27	375	30.5	239	4	US-10-329-624-5239

28	372.5	30.3	239	5	US-10-952-915-44	Sequence 44, Appl
29	372	30.2	232	3	US-09-738-626-3954	Sequence 3954, Ap
30	372	30.2	232	6	US-11-006-098-136	Sequence 136, App
31	371	30.2	235	5	US-10-501-282-962	Sequence 962, App
32	365.5	29.7	244	4	US-10-282-122A-50681	Sequence 50681, A
33	365	29.7	226	4	US-10-282-122A-53860	Sequence 53860, A
34	365	29.7	234	3	US-09-815-242-13265	Sequence 13265, A
35	365	29.7	234	3	US-09-815-242-13558	Sequence 13558, A
36	365	29.7	234	5	US-10-472-928-2464	Sequence 2464, Ap
37	365	29.7	251	4	US-10-282-122A-74003	Sequence 74003, A
38	365	29.7	252	5	US-10-617-320-3288	Sequence 3288, Ap
39	362	29.4	246	4	US-10-724-972A-4146	Sequence 4146, Ap
40	361	29.3	230	5	US-10-501-282-960	Sequence 960, App
41	356	28.9	223	4	US-10-156-761-11506	Sequence 11506, A
42	356	28.9	232	4	US-10-282-122A-52443	Sequence 52443, A
43	352.5	28.7	227	4	US-10-282-122A-61935	Sequence 61935, A
44	352.5	28.7	227	4	US-10-282-122A-64389	Sequence 64389, A
45	352.5	28.7	230	4	US-10-282-122A-62590	Sequence 62590, A

ALIGNMENTS

RESULT 1

US-10-613-990A-32
; Sequence 32, Application US/10613990A
; Publication No. US20040180404A1
; GENERAL INFORMATION:
; APPLICANT: Ajinomoto Co., Inc.
; TITLE OF INVENTION: Method for Producing Target Substance by Fermentation
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/613,990A
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: JP 2002-203764
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 32
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-613-990A-32

Query Match	100.0%	Score 1230;	DB 4;	Length 238;
Best Local Similarity	100.0%	Pred. No. 4.2e-116;		
Matches	238;	Conservative	0;	Mismatches 0;
		Indels	0;	Gaps 0;
Qy	1	MQTPHILIVEDLVTRNTLKSI	FEAEGYDVFEATDGAEMHQIL	SEYDINLVIMDINLPKG 60
Db	1	MQTPHILIVEDLVTRNTLKSI	FEAEGYDVFEATDGAEMHQIL	SEYDINLVIMDINLPKG 60
Qy	61	NGLLARELREQANVALMFLTG	RNEVDKIIGLEIGADDYITKPN	PRELTIARNLISR 120
Db	61	NGLLARELREQANVALMFLTG	RNEVDKIIGLEIGADDYITKPN	PRELTIARNLISR 120
Qy	121	TWNLTGVSEERSVESYKFN	GWELDIINSRSLIGPDGEQYKLP	RSSEFRAMLHFCENPGKIQ 180
Db	121	TWNLTGVSEERSVESYKFN	GWELDIINSRSLIGPDGEQYKLP	RSSEFRAMLHFCENPGKIQ 180
Qy	181	SRABLLKMTQRELKPHDRT	VDVTIRIRKHFESTPDTPPEII	ATIHGEGYRFGCDLED 238
Db	181	SRABLLKMTQRELKPHDRT	VDVTIRIRKHFESTPDTPPEII	ATIHGEGYRFGCDLED 238

RESULT 2

US-10-613-990A-20
; Sequence 20, Application US/10613990A
; Publication No. US20040180404A1
; GENERAL INFORMATION:
; APPLICANT: Ajinomoto Co., Inc.
; TITLE OF INVENTION: Method for Producing Target Substance by Fermentation
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/613,990A

```
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: JP 2002-203764
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Pantoea ananatis
US-10-613-990A-20

Query Match          93.1%; Score 1145; DB 4; Length 238;
Best Local Similarity 92.0%; Pred. No. 1.7e-107;
Matches 219; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

QY 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
DB 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60

QY 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRELTIARNLLSR 120
DB 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRELTIARNLLSR 120

QY 121 TNLGTVSEERSV---ESYKFNCGWELDINSRSLIGPDGEQYKLPKRSFRAMLHFCENPG 177
DB 121 TNLGTVSEERSV---ESYKFNCGWELDINSRSLIGPDGEQYKLPKRSFRAMLHFCENPG 177

QY 178 KIQRABELLKQMTGRELKHDPDRTVDVTTIRIRKHFESTPDPPEIIATIHGEGYRFGCDLE 237
DB 178 KIQRABELLKQMTGRELKHDPDRTVDVTTIRIRKHFESTPDPPEIIATIHGEGYRFGCDLE 237

RESULT 3
US-09-815-242-11143
; Sequence 11143, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Judith W.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11143
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-11143

Query Match          78.4%; Score 964.5; DB 3; Length 236;
Best Local Similarity 76.2%; Pred. NO. 3.5e-89;
Matches 183; Conservative 26; Mismatches 24; Indels 7; Gaps 2;

; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: JP 2002-203764
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Pantoea ananatis
US-10-613-990A-20

Query Match          93.1%; Score 1145; DB 4; Length 238;
Best Local Similarity 92.0%; Pred. No. 1.7e-107;
Matches 219; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

QY 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60
DB 1 MQTPHILIVEDLVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKG 60

QY 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRELTIARNLLSR 120
DB 61 NGLLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRELTIARNLLSR 120

QY 121 TNLGTVSEERSV---ESYKFNCGWELDINSRSLIGPDGEQYKLPKRSFRAMLHFCENPG 177
DB 121 TNLGTVSEERSV---ESYKFNCGWELDINSRSLIGPDGEQYKLPKRSFRAMLHFCENPG 177

QY 178 KIQRABELLKQMTGRELKHDPDRTVDVTTIRIRKHFESTPDPPEIIATIHGEGYRFGCDLE 237
DB 178 KIQRABELLKQMTGRELKHDPDRTVDVTTIRIRKHFESTPDPPEIIATIHGEGYRFGCDLE 237

RESULT 4
US-10-282-122A-58341
; Sequence 58341, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58341
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-282-122A-58341

Query Match          78.4%; Score 964.5; DB 4; Length 236;
Best Local Similarity 76.2%; Pred. NO. 3.5e-89;
Matches 183; Conservative 26; Mismatches 24; Indels 7; Gaps 2;
```



```
QY 1 MOTPHILIVEDLVTRNTLKSIFEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPGK 60
Db 1 MTTPKILVVEDEVTRNTLKGIFEAGYDVFEAENGVEWHILANHNINLVVMDINLPGK 60
QY 61 NGLLLARELREQANVALMFLTRDNEVDKILGLEITGADDTYTKPNPRELTIARNLISR 120
Db 61 NGLLLARELRELSPLIFLTRDNEVDKILGLEITGADDTYTKPNPRELTIARNLISR 120
QY 121 TNLGTVSEERSV--ESYKFNGLWELNRSRSLIGPDGEQYKLPSEFRAMLHFCENPG 177
Db 121 AM----PHQEKENTGREFYRNGWKLNDNSHSLTPGQEFKLPSEFRAMLHFCENPG 176
QY 178 KIQSRAELLKQMTGRELKHDPDRVDTVIRIRKHFESTPDTPEIATIHGEGYRCGDL 237
Db 177 KLTREELLKQMTGRELKHDPDRVDTVIRIRKHFEDHPNTPIIWTIHGEGYRCGDIE 236

RESULT 5
US-10-450-763-59111
; Sequence 59111, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyscg, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 59111
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (26)..(69)
; OTHER INFORMATION: Transcriptional regulatory protein C terminal, domain
; OTHER INFORMATION: identified by eMATRIX, accession number PF00486C, p-value=8.560e-
; OTHER INFORMATION: 34, raw score of 20.21
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (105)..(169)
; OTHER INFORMATION: Transcriptional regulatory protein, C ter domain identified
; OTHER INFORMATION: by Pfam, accession name trans_reg_C, E-value=2.4e-24, Pfam score
; OTHER INFORMATION: 85.8
US-10-450-763-59111

Query Match 63.3%; Score 778; DB 5; Length 176;
Best Local Similarity 93.8%; Pred. No. 1.9e-70;
Matches 152; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

QY 70 REQANVALMFLTRDNEVDKILGLEITGADDTYTKPNPRELTIARNLISRNLGTVSE 129
Db 18 QKQANVALMFLTRDNEVDKILGLEITGADDTYTKPNPRELTIARNLISRNLGTVSE 77
QY 130 ERRSVESYKFNGLWELNRSRSLIGPDGEQYKLPSEFRAMLHFCENPGKIQSRAELLKQM 189
Db 78 ERRSVESYKFNGLWELNRSRSLIGPDGEQYKLPSEFRAMLHFCENPGKIQSRAELLKQM 137
QY 190 TGRELKHDPDRVDTVIRIRKHFESTPDTPEIATIHGEGYR 231
Db 138 TGRELKHDPDRVDTVIRIRKHFESTPDTPEIATIHGEGYR 175

RESULT 6
US-10-282-122A-74434
```

```
; Sequence 74434, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 74434
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-282-122A-74434
```

```
Query Match 32.5%; Score 400; DB 4; Length 236;
Best Local Similarity 39.5%; Pred. No. 6.2e-32;
Matches 92; Conservative 45; Mismatches 82; Indels 14; Gaps 4;

QY 6 ILIVDELVTRNTLKSIFEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPGKNGLL 65
Db 4 ILIVDEKPISDIIFKFNLTKEGYDVTAFDGRAVTFEEKPDLLILDMLPELDGLEV 63
QY 66 ARELREQANVALMFLTRDNEVDKILGLEITGADDTYTKPNPRELTIARNLISRNLG 125
Db 64 AKERKTSHPVPIIMLSAKDSFQKVIGLEITGADDTYTKPNPRELTIARNLISRNLG 123
QY 126 T-VSEERRSVESYKFNGLWELNRSRSLIGPD-----GEQYKLPSEFRAMLHFCENPG 177
Db 124 TAVAEANSSGTQ-----ELTIGNLQIL-PDAFVAKKHQGEVELTHREFFLLHLANHWG 177
QY 178 KIQSRAELLKQMTGRELKHDPDRVDTVIRIRKHFESTPDTPEIATIHGEGY 230
Db 178 QVMTREHLLLEIVMGYDVGDRVTDVTVIRIRKIEKIEDTPSRPEVILTTRGVGY 230

RESULT 7
US-10-282-122A-60754
; Sequence 60754, Application US/10282122A
; Publication No. US20040029129A1
```

; GENERAL INFORMATION:

; APPLICANT: Wang, Lianguo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 60754

; LENGTH: 237

; TYPE: PRT

; ORGANISM: Listeria monocytogenes

US-10-282-122A-60754

Query Match 31.7%; Score 389.5; DB 4; Length 237;
Best Local Similarity 36.0%; Pred. No. 7.3e-31;
Matches 86; Conservative 53; Mismatches 81; Indels 19; Gaps 4;

QY 1 MQTPHILIVEDLVTRNTLKSIFBAGYDVDFEATDGAEMHQILSEYDINLVIMDINLPKG 60
DB 1 MAEKKILVDDDEKPIADILKFNLEKEGFEIWAHDGDEAIEKATBEQDPMVLLDINLPKG 60
QY 61 NGLLARELRQANVALMFLTGRDNEVDKILGLEIGADDDYITKPNPRELTIRARNLLSR 120
DB 61 DGLVCEVRKVKYDPIINVTAKDSEIDKVLGLEIGADDDYITKPNPRELTIRARNLLSR 120
QY 121 --TMNLGTVSERRSVESYKFNKGWELDINSRLIGPD-----GEQVKLPRSEFRAMLH 171
DB 121 HSQVSSSTAEBENS-----BLEIGS-LIHPDAYVASKRGTELTHTREPELLHY 170
QY 172 FCENFGKIQSRAELKKMTGRELKPHDRTVDVTIRIRKHFESTDPTPEIATIHGEY 230
DB 171 LAKHMGQVMTREHLLQTVGWGYDFGVDVTVVTRRLREKIEDNPNFPAVLVTRRGVY 229

RESULT 8

US-10-282-122A-46382

; Sequence 46382, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Lianguo

; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 46382

; LENGTH: 235

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-10-282-122A-46382

Query Match 31.4%; Score 386.5; DB 4; Length 235;
Best Local Similarity 35.1%; Pred. No. 1.5e-30;
Matches 85; Conservative 46; Mismatches 84; Indels 27; Gaps 2;

QY 1 MQTPHILIVEDLVTRNTLKSIFBAGYDVDFEATDGAEMHQILSEYDINLVIMDINLPKG 60
DB 1 MWGKKILVDDDEKPIADILKFNLEKEGFEIWAHDGDEAIEKATBEQDPMVLLDINLPKG 60
QY 61 NGLLARELRQANVALMFLTGRDNEVDKILGLEIGADDDYITKPNPRELTIRARNLLSR 120
DB 61 DGLVCEVRKIRKSEMPIIMLTAKDSEIDKVLGLEIGADDDYITKPFSTRELLARVANLRR 120
QY 121 TMNLGTVSERRSVESYKFNKGWELDINSRLIGP-----DGEQVKLPRSEFRA 168
DB 121 HQQGAARKEE-----NTEMWIGPIVINPNAYSVTYKRENIELTHREPEL 165
QY 169 MLHFCENFGKIQSRAELKKMTGRELKPHDRTVDVTIRIRKHFESTDPTPEIATIHGE 228
DB 166 LHYLAKHLGQVMTREHLLQTVGWGYDFGVDVTVVTRRLREKIEDNFSHPFLIVTRGV 225
QY 229 GY 230
DB 226 GY 227

RESULT 9

US-10-282-122A-72172

; Sequence 72172, Application US/10282122A

; Publication No. US20040029129A1

RESULT 10
US-09-815-242-4913
; Sequence 4913, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert

```

US-001-1
US-09-815-242-10748
; Sequence 10748, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zysekind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, Howard
; TITLE OF INVENTION: Identification of
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELTRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,
; CURRENT FILING DATE: 2001-03-21

```

```

RESULT 13
US-10-282-122A-71628
  Sequence 71628, Application US/10282122A
  Publication No. US20040029129A1
  GENERAL INFORMATION:
  APPLICANT: Wang, Liangsu
  APPLICANT: Zamudio, Carlos
  APPLICANT: Malone, Cheryl
  APPLICANT: Haselbeck, Robert
  APPLICANT: Orlsen, Kari
  APPLICANT: Zyskind, Judith
  APPLICANT: Wall, Daniel
  APPLICANT: Trawick, John
  APPLICANT: Carr, Grant
  APPLICANT: Yamamoto, Robert
  APPLICANT: Forsyth, R.
  APPLICANT: Xu, H.
  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
  FILE REFERENCE: ELITRA.034A
  CURRENT APPLICATION NUMBER: US/10/282,122A
  CURRENT FILING DATE: 2003-02-20
  PRIOR APPLICATION NUMBER: 60/191,078
  PRIOR FILING DATE: 2000-03-21
  PRIOR APPLICATION NUMBER: 60/206,848
  PRIOR FILING DATE: 2000-05-23
  PRIOR APPLICATION NUMBER: 60/207,727
  PRIOR FILING DATE: 2000-05-26
  PRIOR APPLICATION NUMBER: 60/230,335
  PRIOR FILING DATE: 2000-09-06
  PRIOR APPLICATION NUMBER: 60/230,347
  PRIOR FILING DATE: 2000-09-09
  PRIOR APPLICATION NUMBER: 60/242,578

```

```
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71628
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-71628

Query Match      30.6%; Score 376; DB 4; Length 238;
Best Local Similarity 35.6%; Pred. No. 1.7e-29;
Matches 83; Conservative 52; Mismatches 80; Indels 18; Gaps 3;

QY 6 ILIVDELVTRTLKSIPEAGSYDVFEATDCAEMHQILSEYDINLVINDINLPKNGLL 65
Db 10 VVVVDDEKPIADILFNLUKKEGYDYVCAYDGNDAVDLIYSEPDVILDLMLPGRDGMEV 69
QY 66 ARELREQANVALMFTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARNLLSRMTNLG 125
Db 70 CREVKKEMPIIMLTAKDSEIDKVLGLELGADDYVTKPFSTRELIAVRKANLR----- 124
QY 126 TVSERRSVESYKFGNWELDINSRL-IGPD-----GEQYKLPSPFRAMLHFCENPG 177
Db 125 -----HYSQPAQEVSGTTNEITIKDIVIPDAYSIKRGEDIELTHREBFELFHYLSKMG 179
QY 178 KIQSRAELLKMTGRELPHDRTVDVTRIRKHPSTPDPEIATIHGEGY 230
Db 180 QVMTREHLLQTVMGYDFGVDVTRVDRIRLRKEIDDDPSHPHYIVTRRGVGY 232

RESULT 14
US-10-282-122A-57574
; Sequence 57574, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57574
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (18)-(18)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (20)-(20)
; OTHER INFORMATION: X=any amino acid
US-10-282-122A-57574

Query Match      30.6%; Score 376; DB 4; Length 261;
Best Local Similarity 35.0%; Pred. No. 2e-29;
Matches 79; Conservative 55; Mismatches 88; Indels 4; Gaps 2;

QY 6 ILIVDELVTRTLKSIPEAGSYDVFEATDCAEMHQILSEYDINLVINDINLPKNGLL 65
Db 31 VLVVDDEKPIADILFNLUKKEGYDYVCAYDGNDAVDLIYSEPDVILDLMLPGRDGMEV 90
QY 66 ARELREQANVALMFTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARNLLSRMTNLG 125
Db 91 AREVKTYDMPIMLTAKDSEIDKVLGLELGADDYVTKPFSTRELIAVRKANLR----- 147
QY 126 TVSERRSVESYKFGNWELDINSRS-LIGPDGEQYKLPSPFRAMLHFCENPGKIQSRAE 184
Db 148 TAAKEPEEAPAEALTIGDLTIHPEAYMVTKRGETIELTHREBFELFYLAKHLGQVMTREH 207
QY 185 LLKMTGRELPHDRTVDVTRIRKHPSTPDPEIATIHGEGY 230
Db 208 LLQTVMGYDFGVDVTRVDRIRLRKEIDDDPSHPHYIVTRRGVGY 253

RESULT 15
US-09-815-242-5676
; Sequence 5676, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2006, 02:55:25 ; Search time 17 Seconds
(without alignments)
183.716 Million cell updates/sec

Title: US-10-613-990A-32
Perfect score: 1230
Sequence: 1 MPTPHILIVEDLVTRNLK.....PEIIATHGEGYRFGDLED 238

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 97014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New.*
1: /cgm2_6/ptodata/1/pubpaa/US08_NEW_PUB.pap.*
2: /cgm2_6/ptodata/1/pubpaa/US06_NEW_PUB.pap.*
3: /cgm2_6/ptodata/1/pubpaa/US07_NEW_PUB.pap.*
4: /cgm2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pap.*
5: /cgm2_6/ptodata/1/pubpaa/US05_NEW_PUB.pap.*
6: /cgm2_6/ptodata/1/pubpaa/US10_NEW_PUB.pap.*
7: /cgm2_6/ptodata/1/pubpaa/US11_NEW_PUB.pap.*
8: /cgm2_6/ptodata/1/pubpaa/US60_NEW_PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	375	30.5	233	6	US-10-793-626-480
2	285	23.2	238	7	US-11-074-176-148
3	274	22.3	225	6	US-10-467-657-2276
4	262	21.3	224	6	US-10-793-626-3266
5	204	16.6	166	6	US-10-793-626-1424
6	204	16.6	166	6	US-10-793-626-2788
7	198.5	16.1	185	6	US-10-873-528-54
8	175	14.2	122	7	US-11-098-686-11387
9	154	12.5	188	6	US-10-467-657-1982
10	146.5	11.9	425	6	US-10-467-657-8052
11	139	11.3	465	7	US-11-098-686-10932
12	131.5	10.7	141	6	US-10-793-626-686
13	131.5	10.7	141	6	US-10-793-626-2234
14	130.5	10.6	471	7	US-11-098-686-11324
15	128.5	10.4	457	7	US-11-098-686-11122
16	124	10.1	251	6	US-10-793-626-238
17	124	10.1	251	6	US-10-793-626-722
18	122	9.9	126	7	US-11-098-686-10771
19	121	9.8	218	6	US-10-467-657-4822
20	121	9.8	2087	7	US-11-075-185-28
21	114.5	9.3	461	7	US-11-075-185-25
22	114	9.3	255	7	US-11-075-185-20
23	113	9.2	316	7	US-11-098-686-10974
24	109.5	8.9	740	7	US-11-140-625-12
25	109	8.9	218	6	US-10-793-626-108

26	108	8.8	738	7	US-11-140-625-11	Sequence 11, Appl
27	105	8.5	132	6	US-10-667-295-147	Sequence 147, Appl
28	104	8.5	141	6	US-10-667-295-25	Sequence 25, Appl
29	104	8.5	166	6	US-10-667-295-24	Sequence 24, Appl
30	104	8.5	170	6	US-10-667-295-23	Sequence 23, Appl
31	96	7.8	754	7	US-11-140-625-10	Sequence 10, Appl
32	90.5	7.4	212	6	US-10-454-437-298	Sequence 298, Appl
33	88.5	7.2	319	7	US-11-098-686-11362	Sequence 11362, A
34	87.5	7.1	884	7	US-11-098-686-11269	Sequence 11269, A
35	85	6.9	142	6	US-10-667-295-74	Sequence 74, Appl
36	85	6.9	3056	7	US-11-109-156-20	Sequence 20, Appl
37	82	6.7	207	6	US-10-467-657-2592	Sequence 2592, Ap
38	82	6.7	210	6	US-10-454-437-302	Sequence 302, App
39	82	6.7	363	6	US-10-995-561-541	Sequence 541, App
40	82	6.7	725	6	US-10-467-657-1536	Sequence 1536, Ap
41	81.5	6.6	208	6	US-10-793-626-782	Sequence 782, App
42	81	6.6	131	6	US-10-667-295-34	Sequence 34, Appl
43	81	6.6	136	6	US-10-667-295-33	Sequence 33, Appl
44	81	6.6	341	6	US-10-515-481-3	Sequence 3, Appl
45	79.5	6.5	738	7	US-11-140-625-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-10-793-626-480
; Sequence 480, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 480
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-480

Query Match	30.5%	Score	375;	DB	6;	Length	233;
Best Local Similarity	35.6%	Pred. No.	1.5e-26;				
Matches	83;	Conservative	52;	Mismatches	80;	Indels	18;
						Gaps	3;
QY	6	ILIVDELVTRNTLKSIFEAEGYDVFEATDGAEMHQILISEYDINLVINDINLPGKNGLL	65				
Db	5	VVVVDDEKPIADILLEFLNKKGGYDVYCAVDGNDVDAVLIESEEPDIVLDDIMLPGRDGM	64				
QY	66	ARELRQANVAMFLTGTGRDNEVDKILGLEIGADDYITKFPNPRELTIRARLLSRMTNLG	125				
Db	65	CRVRKKYEMFIIMLTAKDSEIDKVLGLGADDDYVTKFSTRELIARVKNLR-----	119				
QY	126	TVSERRSVESYKFGWELDINSRL-IGPD-----GEQYKLPFRSEFRAMLHFCENPG	177				
Db	120	-----HYSQPAQEVSGATNEITKIDVIYPDAYSIKKRGEDIETHREFELFYLKHWG	174				
QY	178	KIQSRALLKQMTGRELKPHDRVTVDVIRIRKHKFESTPDTPPEIIATHHGEGY	230				
Db	175	QVMTREHLLQTVMGVDYFGDVRTVDVTIRRLREKIEDDPSHPHVIIVTRRGVY	227				

RESULT 2
US-11-074-176-148
; Sequence 148, Application US/11074176
; Publication No. US20050250135A1

```

; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Perli, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-148

Query Match 23.2%; Score 285; DB 7; Length 238;
Best Local Similarity 30.1%; Pred. No. 1.7e-18;
Matches 69; Conservative 47; Mismatches 104; Indels 10; Gaps 3;

QY 6 ILIIVDELVTRNLKSIPEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPKNGLL 65
DB 4 ILIIEDEKNLARFVLELQHENYETVVENNGRKGDDALLAQDFAILLDMLPDLNGLEI 63

QY 66 ARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRELTIARNLLSRT--- 121
DB 64 ARRVQKVTTPILMTARDSDVDRVSGLDHGADDIVRPFATEELLARLRAVLRVKIEK 123

QY 122 -MNLGTVSEERSVSYKPNGWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPGKIQ 180
DB 124 DASKVTVAQK---IVKFKOLTETANRIVHRGDKAIDLTKREYNLLMTLIENKNVV 179

QY 181 SPAELKKMTGRELKHDPDRTVDVTIRIRKHFESTPDTPEIATIHGEY 230
DB 180 SRDQLLNKIGWPEISNIETNVVEVRYLRNKIDA-PGQPSYIKTVRGTYG 228

RESULT 3
US-10-467-657-2276
; Sequence 2276, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 2276
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2276

Query Match 22.3%; Score 274; DB 6; Length 225;
Best Local Similarity 30.1%; Pred. No. 1.5e-17;
Matches 71; Conservative 47; Mismatches 92; Indels 26; Gaps 6;

QY 6 ILIIVDELVTRNLKSIPEAGYDVFEATDG-AEMHQILS-EYDINLVIMDINLPKNGL 63
DB 4 VLLVDDALLTELLTEYLSAEGLNRSVPDGEAGVQEILSGQYDV--VVLDSPMPKNGL 61

; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1424

US-10-793-626-1424
; Sequence 1424, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1424

QY 64 LLARELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRELTIARNLLSRTMN 123
DB 62 DVLKNVRARSTVPIIMLTAKGDDIDRIIGLEWGADDYVVKPCTPRELLARINAILURRAQH 121

QY 124 LGTVSEERSVSYKPNGWELDINSRLIGPDGBQ-----YKLPSEFRAMLHFCENP 176
DB 122 SGEQNNAFNSISV-----SDVLYPAKQASVKDMPLELTSTEFNLLVLMRHA 170

QY 177 KQISRAELKKMTGRELKHDPDRTVDVTIRIRKHFESTPDTPEIATIHGEYRFP 232
DB 171 QGVVSKETLSVEALDRKLAKEDRSIDVHISIRHKLGD-----SLIQTVRGGLGYLF 222

RESULT 4
US-10-793-626-3266
; Sequence 3266, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3266
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-3266

Query Match 21.3%; Score 262; DB 6; Length 224;
Best Local Similarity 28.9%; Pred. No. 1.8e-16;
Matches 67; Conservative 45; Mismatches 96; Indels 24; Gaps 5;

QY 7 LIVDELVTRNLKSIPEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPKNGLLLA 66
DB 5 LIVDDKKLQVSSHLERESIQTHFTTSGEASLDLFLENKQVDIAIVDIMNMGMDGFELC 64

QY 67 RELREQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRELTIARNLLSR----- 120
DB 65 QTLKDDYHIPVIMLTARDALSDKERAFLSGTDYVTKPFVKELLFRIKAVLRRYQINAD 124

QY 121 -TMNLGTVSEERSVSYKPNGWELDINSRLIGPDGEQYKLPSEFRAMLHFCENPGKI 179
DB 125 NELQLGNLIIN-----QSY----MEITVGSKTM-----NLPNKEFQLFLASNPKHI 168

QY 180 QSRAECLKKMTGRELKHDPDRTVDVTIRIRKHFESTPDTPEIATIHGEYR 231
DB 169 FTRDDIIEKINGFDYEGDDRTVDVHIKELRQLRSLKLSVSV-IQTVRGQYR 219

RESULT 5
US-10-793-626-1424
; Sequence 1424, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1424
```



```
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1424

Query Match      16.6%; Score 204; DB 6; Length 166;
Best Local Similarity 28.9%; Pred. No. 1.8e-11;
Matches 52; Conservative 33; Mismatches 71; Indels 24; Gaps 4;

Qy 57 LPKNGKLLARELRQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARN 116
Db 2 LPNGLGICQIRQKTTPIIIITAKSETYDKVAGLDYGADDYIVKPFIEELLARIRA 61

Qy 117 LLSRTMNLGTSEERSRSVESYKFNCGWELDINSRSL-----IGPDGEQYKLPSEFRAML 170
Db 62 VLRR-----QPKDQV-----LDINGIIIDKDAFKVTYNGHQLEITKTEYDILLY 104

Qy 171 HFCENPGKIQRARELLKMTGRELKPHDRTVDVTIRIRKHFESPTDTPPEIATIHGEGY 230
Db 105 VLAENRNHVQREQLDHWGYNSEVETNVVDVYIRYLNRKLPKPP-NKEKSIETVRGVGY 163

RESULT 6
US-10-793-626-2788
; Sequence 2788, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2788
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-2788

Query Match      16.6%; Score 204; DB 6; Length 166;
Best Local Similarity 28.9%; Pred. No. 1.8e-11;
Matches 52; Conservative 33; Mismatches 71; Indels 24; Gaps 4;

Qy 57 LPKNGKLLARELRQANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARN 116
Db 2 LPNGLGICQIRQKTTPIIIITAKSETYDKVAGLDYGADDYIVKPFIEELLARIRA 61

Qy 117 LLSRTMNLGTSEERSRSVESYKFNCGWELDINSRSL-----IGPDGEQYKLPSEFRAML 170
Db 62 VLRR-----QPKDQV-----LDINGIIIDKDAFKVTYNGHQLEITKTEYDILLY 104

Qy 171 HFCENPGKIQRARELLKMTGRELKPHDRTVDVTIRIRKHFESPTDTPPEIATIHGEGY 230
Db 105 VLAENRNHVQREQLDHWGYNSEVETNVVDVYIRYLNRKLPKPP-NKEKSIETVRGVGY 163

RESULT 7
US-10-873-528-54
; Sequence 54, Application US/10873528
; Publication No. US2005027681A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
```

```
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/10/873,528
; CURRENT FILING DATE: 2004-06-23
; PRIOR APPLICATION NUMBER: US/09/769,787
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 54
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-873-528-54

Query Match      16.1%; Score 198.5; DB 6; Length 185;
Best Local Similarity 30.3%; Pred. No. 6.4e-11;
Matches 47; Conservative 28; Mismatches 63; Indels 17; Gaps 2;

Qy 3 TPHILIVDELVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKNG 62
Db 2 TKQVLLVDDESHILKLLDYHLKSGFSTQLVNTRKALALAEPTFPDFILLDMLPQLDGG 61

Qy 63 LLLARELREQ--ANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARNLLSRT 121
Db 62 MEVCKRLRAKGVKTPIMVSAKSDEFKVLALALEIGADDYITKPFSPRELLARVAVLRRT 121

Qy 122 MN-----LGTVSEERSRSVESYKFN 140
Db 122 KGEQEGDSDNIADDSWLFGLTKVYPERHEVYKAN 156

RESULT 8
US-11-098-686-11387
; Sequence 11387, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11387
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-11387

Query Match      14.2%; Score 175; DB 7; Length 122;
Best Local Similarity 31.4%; Pred. No. 4.5e-09;
Matches 37; Conservative 29; Mismatches 48; Indels 4; Gaps 2;

Qy 5 HILVDELVTRNTLKSIFEAEGYDVFEATDGAEMHQILSEYDINLVIMDINLPKNGKL 63
Db 4 HILVDDSKTVRNLMSPFVKSQGVQVTTAEGLGLEKLYSVKVDKVLIVSINWPRMDGF 63

Qy 64 LLLARELREQ--ANVALMFLTGRDNEVDKILGLEIGADDYITKPFNPRLTIRARNLL 118
Db 64 TFIKSVREQDLYRNIPILISTEKEDEDIKGMKFGANLYMVKPAQPDPTLVNRNIKMLL 121

RESULT 9
US-10-467-657-1982
```

; Sequence 1982, Application US/10467657
 ; Publication No. US20050260581A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SpA
 ; APPLICANT: FONTANA Maria Rita
 ; APPLICANT: PIZZA Mariagrazia
 ; APPLICANT: MASIGNANI Vega
 ; APPLICANT: MONACI Elisabetta
 ; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/10/467,657
 ; CURRENT FILING DATE: 2003-08-11
 ; PRIOR APPLICATION NUMBER: GB-0103424.8
 ; PRIOR FILING DATE: 2001-02-12
 ; NUMBER OF SEQ ID NOS: 9218
 ; SOFTWARE: SeqWin99, version 1.04
 ; SEQ ID NO 1982
 ; LENGTH: 188
 ; TYPE: PRT
 ; ORGANISM: Neisseria gonorrhoeae
 US-10-467-657-1982

 Query Match 12.5%; Score 154; DB 6; Length 188;
 Best Local Similarity 29.5%; Pred. No. 6.2e-07;
 Matches 38; Conservative 26; Mismatches 45; Indels 20; Gaps 2;

 QY 27 GYDVFEATDGAEMHQILSEYDINLVIMDINLPGKGLLLARELREQ-ANVALMELTGRDN 85
 DB 9 GYADVWVKNQAAAAAQPVDLMLDGLPGRGLDVLSEIRAAAGCTVPVLIVTARD 68

 QY 86 EVDKILGLEIGADDYITKFPNPRELTIARNLLSR-----TMNLGT 126
 DB 69 LYSRLNGLDGGADDYIVKFPDMAEPKARMAVLRRGSGOACLSNGALSINPATYQVEI 128

 QY 127 VSERRSVE 135
 DB 129 IAEGRGGIE 137

 RESULT 10
 US-10-467-657-8052
 ; Sequence 8052, Application US/10467657
 ; Publication No. US20050260581A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SpA
 ; APPLICANT: FONTANA Maria Rita
 ; APPLICANT: PIZZA Mariagrazia
 ; APPLICANT: MASIGNANI Vega
 ; APPLICANT: MONACI Elisabetta
 ; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/10/467,657
 ; CURRENT FILING DATE: 2003-08-11
 ; PRIOR APPLICATION NUMBER: GB-0103424.8
 ; PRIOR FILING DATE: 2001-02-12
 ; NUMBER OF SEQ ID NOS: 9218
 ; SOFTWARE: SeqWin99, version 1.04
 ; SEQ ID NO 8052
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Neisseria gonorrhoeae
 US-10-467-657-8052

 Query Match 11.9%; Score 146.5; DB 6; Length 425;
 Best Local Similarity 26.5%; Pred. No. 9e-06;
 Matches 40; Conservative 34; Mismatches 62; Indels 15; Gaps 2;

 QY 1 MQPPIILVDELVTRNTLKSIFEAGYDVFEATDGAEMHQILSEYDINLVIMDINLPGK 60
 DB 1 MRSSDILVDEVGIRDLSEILQDEGYVALAENAEARKURHQARFAMVLLDIWMPDC 60

 QY 61 NGLLALRELE--QANVALMFLTGRDNEVDKILGLEIGADDYITKFPNPRELTIARNLL 118

Query Match 10.7%; Score 131.5; DB 6; Length 141;
Best Local Similarity 27.2%; Pred.No. 4.2e-05;
Matches 40; Conservative 29; Mismatches 65; Indels 13; Gaps 5;

87 VDKILGLEICADDYITKPFNPRLTIRARNLLSRTNMNLGTVSVSRSSVESYKFNGWE--- 143
Db :|::: |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
1 MDQVMSWELGADDMQMFPFNTNVIKLQAIYRVVEFGV--BEKRTL-----SWDAT 52

144 LDINRSLSIGPDGEQYLKPSEFRAMLHFCENFGPKIQSAEALLKMTGTRELKPHDRTVDV 203
Qy |:|:|: |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~:~:
Db |:|:|: |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~:~:
53 VDL-SKDSIQDDKTIFLSKTEMLLEMLINKRNIQIVTRDTLTALWDDEAFVSNDTLTV 111

204 TIRIRKHFEPTDPETPEIATIHGGY 230
Qy :||:||: |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~:~:
Db :||:||: |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~:~:
112 NVNLRKKL-SEIDMDSAIETKVGKY 137

RESULT 13
US-10-793-626-2234
; Sequence 2234, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2234
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-2234

Query Match 10.7%; Score 131.5; DB 6; Length 141;
Best Local Similarity 27.2%; Pred.No. 4.2e-05;
Matches 40; Conservative 29; Mismatches 65; Indels 13; Gaps 5;

87 VDKILGLEICADDYITKPFNPRLTIRARNLLSRTNMNLGTVSVSRSSVESYKFNGWE--- 143
Qy :|::: |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~:~:
Db 1 MDQVMSWELGADDMQMFPFNTNVIKLQAIYRVVEFGV--BEKRTL-----SWDAT 52

144 LDINRSLSIGPDGEQYLKPSEFRAMLHFCENFGPKIQSAEALLKMTGTRELKPHDRTVDV 203
Qy |:|:|: |:|:|:|:|:|:|:|:|:|:|:~::~:~:
Db |:|:~::~:~:
53 VDL-SKDSIQDDKTIFLSKTEMLLEMLINKRNIQIVTRDTLTALWDDEAFVSNDTLTV 111

204 TIRIRKHFEPTDPETPEIATIHGGY 230
Qy :||:||: |||:~::~:~:
Db :||:||: |||:~::~:~:
112 NVNLRKKL-SEIDMDSAIETKVGKY 137

RESULT 14
US-11-098-686-11324
; Sequence 11324, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04

THIS PAGE BLANK (USPTO)